JOURNAL OF BACTERIOLOGY, Nov. 2010, p. 6001–6016 Vol. 192, No. 22
0021-9193/10/$12.00 doi:10.1128/JB.00778-10


# *Legionella pneumophila* Strain 130b Possesses a Unique Combination of Type IV Secretion Systems and Novel Dot/Icm Secretion System Effector Proteins▽†

Gunnar N. Schroeder,[1]‡ Nicola K. Petty,[2]‡ Aurélie Mousnier,[1] Clare R. Harding,[1] Adam J. Vogrin,[4] Bryan Wee,[2] Norman K. Fry,[3] Timothy G. Harrison,[3] Hayley J. Newton,[4] Nicholas R. Thomson,[5] Scott A. Beatson,[2] Gordon Dougan,[5] Elizabeth L. Hartland,[4] and Gad Frankel[1]*

*Centre for Molecular Microbiology and Infection, Division of Cell and Molecular Biology, Imperial College London, London SW7 2AZ, United Kingdom[1]; Centre for Infectious Disease Research, School of Chemistry and Molecular Biosciences, University of Queensland, Brisbane, Queensland 4072, Australia[2]; Respiratory and Systemic Infection Laboratory, Health Protection Agency Centre for Infections, London, NW9 5EQ, United Kingdom[3]; Department of Microbiology and Immunology, University of Melbourne, Victoria 3010, Australia[4]; and Pathogen Genomics, The Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge, CB10 1SA, United Kingdom[5]*

Received 2 July 2010/Accepted 26 August 2010

***Legionella pneumophila* is a ubiquitous inhabitant of environmental water reservoirs. The bacteria infect a wide variety of protozoa and, after accidental inhalation, human alveolar macrophages, which can lead to severe pneumonia. The capability to thrive in phagocytic hosts is dependent on the Dot/Icm type IV secretion system (T4SS), which translocates multiple effector proteins into the host cell. In this study, we determined the draft genome sequence of *L. pneumophila* strain 130b (Wadsworth). We found that the 130b genome encodes a unique set of T4SSs, namely, the Dot/Icm T4SS, a Trb-1-like T4SS, and two Lvh T4SS gene clusters. Sequence analysis substantiated that a core set of 107 Dot/Icm T4SS effectors was conserved among the sequenced *L. pneumophila* strains Philadelphia-1, Lens, Paris, Corby, Alcoy, and 130b. We also identified new effector candidates and validated the translocation of 10 novel Dot/Icm T4SS effectors that are not present in *L. pneumophila* strain Philadelphia-1. We examined the prevalence of the new effector genes among 87 environmental and clinical *L. pneumophila* isolates. Five of the new effectors were identified in 34 to 62% of the isolates, while less than 15% of the strains tested positive for the other five genes. Collectively, our data show that the core set of conserved Dot/Icm T4SS effector proteins is supplemented by a variable repertoire of accessory effectors that may partly account for differences in the virulences and prevalences of particular *L. pneumophila* strains.**

Many bacterial pathogens use specialized protein secretion systems to deliver into host cells virulence effector proteins that interfere with the antimicrobial responses of the host and facilitate the survival of the pathogen (5, 10, 22, 76). The components of these secretion systems are highly conserved. Comparative bioinformatic analysis of pathogen genomes revealed an ever-increasing number of proteins that are likely to be translocated virulence effectors. Only a few effectors have been characterized, and their biochemical functions are unknown, yet the identification of translocated effector proteins and their mechanism of action is fundamental to understanding the pathogenesis of many bacterial infections.

*Legionella pneumophila* is the etiological agent of Legionnaires' disease, which is an acute form of pneumonia (34, 66). *L. pneumophila* serogroup 1 accounts for more than 90% of all cases worldwide. Although *L. pneumophila* is an environmental organism, its ability to survive and replicate in amoebae, such as *Acanthamoeba castellanii*, has equipped the organism with the capacity to replicate in human cells (45, 58, 68, 80). Following the inhalation of aerosols containing *L. pneumophila* into the human lung, the bacteria promote their uptake by alveolar macrophages and epithelial cells (21, 44, 71), where they replicate within an intracellular vacuole that avoids fusion with the endocytic pathway (46, 47). *L. pneumophila* evades endosome fusion by establishing a replicative vacuole that shares many characteristics with the endoplasmic reticulum (ER) (48, 53, 89). The formation of the unique *Legionella*-containing vacuole (LCV) requires the Dot (defective in organelle trafficking)/Icm (intracellular multiplication) type IV secretion system (T4SS) (85, 91).

Type IV secretion systems are versatile multiprotein complexes that can transport DNA and proteins to recipient bacteria or host cells (19, 36). Based on structural and organizational similarity, three main T4SS classes have been distinguished: T4SSA, T4SSB, and genomic island-associated T4SS (GI-T4SS) (3, 51). The genetic organization and components of T4SSA have high similarity to the classical VirB4/VirD4

* Corresponding author. Mailing address: Flowers Building, Imperial College, London SW7 2AZ, United Kingdom. Phone: 44 20 7594 5253. Fax: 44 20 7594 3069. E-mail: g.frankel@imperial.ac.uk.

‡ G.N.S. and N.K.P. contributed equally to this work.

† Supplemental material for this article may be found at http://jb.asm.org/.

▽ Published ahead of print on 10 September 2010.

transfer DNA (T-DNA) transfer system of *Agrobacterium tumefaciens* (3). In the sequenced *L. pneumophila* strains, three distinct T4SSAs with different prevalences among strains have been described: Lvh, Trb-1, and Trb-2 (37, 83, 86). The Lvh (*Legionella* vir homologues) T4SSA is not required for intracellular bacterial replication in macrophages and amoebae but seems to contribute to infection at lower temperatures and inclusion in *Acanthamoeba castellanii* cysts (6, 78, 86).

The Dot/Icm T4SSB secretes and translocates multiple bacterial effector proteins into the vacuolar membrane and cytosol of the host cell (31, 70). The functions of the great majority of these proteins are unknown. Several effectors have similarity to eukaryotic proteins or carry eukaryotic motifs (7, 16, 25). They are predicted to allow *L. pneumophila* to manipulate host cell processes by functional mimicry (31, 70). Many of the effectors have paralogues or belong to related protein families that are likely to have overlapping functions.

Comparative analysis of the recent *L. pneumophila* genome sequences has revealed their diversity and plasticity (16, 18, 88). This plasticity enables the bacterium to acquire new genetic factors, including new effector proteins that enhance bacterial replication and survival in eukaryotic cells. This has resulted in a diverse species in which 7 to 11% of the genes in each *L. pneumophila* isolate are strain specific (38). Some of the diversity occurs among genes encoding Dot/Icm effectors, including those within the same family. For example some ankyrin repeat and F-box effector genes are highly conserved, while others vary considerably between *L. pneumophila* isolates (16, 41, 62, 73, 75). Even though it is not experimentally proven, the subsequent selection of Dot/Icm effectors among different *L. pneumophila* isolates might reflect their usefulness in host-pathogen interactions, whereby different effector repertoires are maintained during adaptation to different environmental niches or hosts. This may then translate into differences in virulence during opportunistic infection.

In this study, we sequenced the genome of *L. pneumophila* serogroup 1 strain 130b (ATCC BAA-74, also known as Wadsworth or AA100) (29, 30) and analyzed the sequence for T4SSs and novel Dot/Icm effectors. This analysis established that the strain encodes a unique combination of T4SSs and a set of Dot/Icm effectors that had not been described previously but that are present in a range of clinical and environmental *L. pneumophila* isolates. The new effectors represent the latest members of an ever-growing list of T4SS substrates and presumably reflect the great capacity of *L. pneumophila* for adaptation to a variety of hosts.

## MATERIALS AND METHODS

**Bacterial strains and sequencing.** The sequenced *L. pneumophila* serogroup 1 strain 130b is a clinical isolate from the Wadsworth Veterans Administration Hospital, Los Angeles, CA (29, 30). The *L. pneumophila* ΔDotA strain is a *dotA* insertion mutant (kanamycin resistance) of *L. pneumophila* strain 130b (84).

*L. pneumophila* 130b was obtained from Nick Cianciotto, Northwestern University, and was subjected to minimal passages. High-purity chromosomal DNA was prepared for sequencing by proteinase K digestion, phenol-chloroform extraction, and ethanol precipitation (90). The whole genome of 130b was sequenced using paired-end 454 FLX pyrosequencing and assembled using the 454/Roche Newbler assembly program into 274 contigs (N50 contig size, 35,584 bp) from 248,625 sequence reads with an average read length of 157 bp. The contigs were scaffolded using paired reads, with an average pair distance of 3,028 bp, into 11 scaffolds (N50 scaffold size, 2,421,541 bp). The *lvh* collapsed repeat region was reassembled into two distinct *lvh* regions by first separating reads that mapped to the previously sequenced 130b *lvh* region (accession no. AF410854), differentiating *lvh* reads according to microheterogeneity (single-nucleotide polymorphism [SNP] content), and then using Newbler (with stringent cutoffs) to generate individual *lvh* assemblies that could be unambiguously positioned within a scaffold in the main assembly. Several contig gaps were closed using PCR and Sanger sequencing in conjunction with manual examination of the individual 454 reads in Consed (39). Unassembled contigs of less than 300 bp were removed from the end of the assembly. This gave a total sequence length of 3,473,547 bp assembled into 145 contigs in 4 scaffolds plus 14 small contigs (114 to 3,270 bp) that could not be scaffolded.

The 130b genome was aligned with the other sequenced *L. pneumophila* genomes to aid whole-genome comparisons, making the first gene *dnaA*. An automated annotation was performed on the genome sequence using SUGAR, as previously described (93). Artemis (82) was used to facilitate the manual curation of the sequence and annotation of the effectors and T4SSs.

**Bioinformatic analysis.** Pairwise whole-genome comparisons of the 130b genome with the other sequenced *L. pneumophila* genomes—Lens (accession number CR628337), Philadephila-1 (accession number AE017354), Corby (accession number CP000675), and Paris (accession number CR628336)—and *Legionella longbeachae* NSW150 (accession number FN650140) were performed using BLASTn and visualized using the Artemis Comparison tool (ACT) (13). Genome comparison figures were made using easyfig (http://easyfig.sourceforge.net/), and the circular diagram was produced with DNAplotter (12).

Bioinformatic analysis of domains and motifs of individual effector protein candidates was performed using the Pfam database (Pfam release 24, HMMER3.0 beta 3 [32]), SMART (version 6 [61]), and the NCBI Conserved Domain Database (version 2.18 [65]).

**Phylogenetic analysis.** Thirty genes in the Philadelphia-1 genome were identified from Hidden Markov-Model (HMM) profiles of genes previously identified as conserved across all bacteria (96). Homologues of these genes were then searched for using tBLASTn (2) in the genomes of *L. pneumophila* strains 130b, Corby, Lens, and Paris and *L. longbeachae* NSW150. Each gene set was manually checked to ensure all genes were intact, and four genes (*dnaG*, *pyrG*, *rplF*, and *rpsJ*) that were disrupted by sequencing errors in the unfinished 130b genome were removed from the data set. The remaining 26 single-copy genes (*frr*, *infC*, *nusA*, *pgk*, *rplA*, *rplB*, *rplC*, *rplD*, *rplE*, *rplK*, *rplL*, *rplM*, *rplN*, *rplP*, *rplS*, *rplT*, *rpmA*, *rpoB*, *rpsB*, *rpsC*, *rpsE*, *rpsI*, *rpsK*, *rpsM*, *smpB*, and *tsf*) were concatenated for each genome, corresponding to a total of 18.5 kb, and aligned using CLUSTALW2 (59).

A maximum-likelihood (ML) tree was built using PhyML (40). The general time reversible substitution model with gamma-distributed rate variation was used for most of the individual alignments, as suggested by FindModel (74). The transition/transversion ratio, proportion of invariable sites, and gamma distribution parameter were estimated by PhyML.

**Strain and plasmid construction.** The plasmids pXDC61, pXDC61-C2 (pXDC61 LegC2), and pXDC61 FabI were a kind gift from Xavier Charpentier and Howard Shuman (Columbia University Medical Center, New York, NY) and were described previously (24). All plasmids, primers, and restriction enzymes used to construct the pXDC61-derived expression vectors for β-lactamase (TEM1) fusions of the putative new *L. pneumophila* strain 130b effector proteins are listed in Table 1. All genes were PCR amplified from *L. pneumophila* strain 130b genomic DNA, and the PCR products were digested and ligated into pXDC61. The sequence identities and correct orientation of the inserts were verified by DNA sequencing. The new plasmids were transformed into *L. pneumophila* strain 130b wild type and ΔDotA by electroporation.

**Translocation assay.** The β-lactamase (TEM1) translocation assay for the identification of Icm/Dot T4SS substrates was adapted from the protocol described by de Felipe et al. (24). Raw264.7 macrophages were cultured in a humidified atmosphere of 5% $CO_2$ at 37°C in RPMI 1640 glutamine medium supplemented with 10% fetal calf serum and Glutamax (Invitrogen). To obtain a confluent cell layer, $7.5 \times 10^4$ Raw264.7 macrophages were seeded in 200 μl growth medium per well of a black wall/clear flat-bottom 96-well plate (Becton Dickinson) and cultured overnight, and the medium was replaced with 150 μl fresh growth medium immediately prior to infection. *L. pneumophila* strain 130b wild type or the ΔDotA mutant harboring the pXDC61-derived TEM1 fusion expression plasmids was inoculated in ACES yeast extract (AYE) broth to an optical density at 600 nm ($OD_{600}$) of 0.1 from 3-day-old charcoal-yeast extract (CYE) plates and grown in the presence of 6 μg/ml chloramphenicol and 0.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) at 37°C for 21 h (post-exponential growth phase) (92). The bacteria were diluted in Raw264.7 macrophage growth medium and added to the macrophages at a multiplicity of infection (MOI) of 40, and the infection was synchronized by centrifugation ($900 \times g$; 10 min). After 1 h of incubation at 37°C, 5% $CO_2$, the supernatant was replaced by 100 μl Hanks'

TABLE 1. Plasmids, cloning primers, and restriction sites used in this study

| Plasmid | Description | Primer sequence (5′–3′) | RS[a] | Source |
|---|---|---|---|---|
| pXDC61 | Expression vector for C-terminal TEM1 fusion protein | | | 24 |
| pXDC61-C2 | Expression vector for LegC2-TEM1 fusion protein | | | 24 |
| pXDC61 FabI | Expression vector for FabI-TEM1 fusion protein | | | 24 |
| pICC536 | pXDC61 *lpw_00581*; expression vector for Lpw_00581 (LtpA)-TEM1 fusion protein | AGGGTACCAGCGTAATAATTGATTCCTTAAAAAATTC<br>CGATGGTACCTTAGCGATTTATATTAAATAAAGATCCATG | KpnI<br>KpnI | This study |
| pICC537 | pXDC61 *lpw_02301*; expression vector for Lpw_02301 (LtpB)-TEM1 fusion protein | AGGGTACCGCATCGGGGAAACAAGATACCTCAC<br>CGTCTAGATCAATTATTTAGTCCGAGATTTGAGG | KpnI<br>XbaI | This study |
| pICC538 | pXDC61 *lpw_02381*; expression vector for Lpw_02381 (LtpC)-TEM1 fusion protein | CGCGCGGTACCCACAATAAAATACTTCAGTTAGTCAACTCTGATC<br>CGCGCTCTAGATCAACTACTGTGTAAAAAAACTTTAGGGGTG | KpnI<br>XbaI | This study |
| pICC539 | pXDC61 *lpw_03701*; expression vector for Lpw_03701 (LtpD)-TEM1 fusion protein | CACTGACCCGGGGAACCAAGAAACATTTAAATCCAGAC<br>GTCGCAAAGCTTCTAAATCTTTAAACCAGTGTCGC | XmaI<br>HindIII | This study |
| pICC540 | pXDC61 *lpw_04551*; expression vector for Lpw_04551 (LtpE)-TEM1 fusion protein | TCGCACTGAGGTACCCCAGATCATAGTAAATTTCCATTTA<br>CGTCGCAGGTACCTTAATAAATTATCCTCGCTAGCTTA | KpnI<br>KpnI | This study |
| pICC541 | pXDC61 *lpw_16311*; expression vector for Lpw_16311 (LtpF)-TEM1 fusion protein | TCGCACTGAGGATCCATTACAACAGAAATCCTAAATATAAA<br>GACGTCGCAAAGCTTTTATTTTAGAGACGGAATTTTAGA | BamHI<br>HindIII | This study |
| pICC542 | pXDC61 *lpw_20091*; expression vector for Lpw_20091 (LtpG)-TEM1 fusion protein | TCGCACTGAGGTACCTTTGATTTAGCAAAACTGACTGAAT<br>GACGTCGCATCTAGATCAAAACATCGCAACAAATTCTG | KpnI<br>XbaI | This study |
| pICC543 | pXDC61 *lpw_20341*; expression vector for Lpw_20341 (LtpH)-TEM1 fusion protein | CACTGAGGATCCCCCAGTTGCAAAAATAACCCC<br>GTCGCAAAGCTTTCATAAAATAGAGCAGTAGGACCA | BamHI<br>HindIII | This study |
| pICC544 | pXDC61 *lpw_21901*; expression vector for Lpw_21901-TEM1 fusion protein | GCACTGAGGTACCAACATTATTTTAATTGGTGAAGTACAT<br>ACTGCGCATCTAGATTACATTAAAGCACAGTTATTTTCTG | KpnI<br>XbaI | This study |
| pICC545 | pXDC61 *lpw_25791*; expression vector for Lpw_25791 (LtpI)-TEM1 fusion protein | CGCACTGAGGTACCCCTTTTCAACTTCCTTTTGAAAAA<br>GACGTCGCATCTAGATTACATGAATCTTACAGAATGGCT | KpnI<br>XbaI | This study |
| pICC546 | pXDC61 *lpw_26201*; expression vector for Lpw_26201 (LtpJ)-TEM1 fusion protein | TCGCACTGAGGTACCTATTACCAGCTAAACCTTAGTGAAC<br>GACGTCGCATCTAGATTAACCAGCTGCCAGGCAAG | KpnI<br>XbaI | This study |
| pICC547 | pXDC61 *lpw_28221*; expression vector for Lpw_28221-TEM1 fusion protein | GCACTGAGGTACCAATTATTTACCTCGACGCATAG<br>CGTCGCAGGTACCAATTATTTAGATAGATCAAGAACAA | KpnI<br>KpnI | This study |

[a] RS, restriction site.

buffered salt solution (Gibco) supplemented with 20 mM HEPES and 3 mM probenecid, pH 7.4 (HBSS-HP), and 20 μl freshly prepared CFF2-AM β-lactamase substrate (LiveBLAzer FRET-B/G Loading Kit; Invitrogen) was added. After 1 h 45 min of incubation at room temperature in the dark, the cells were washed four times with HBSS-HP. Fluorescence emission at 450 nm and 520 nm was measured from the bottom using a Fluostar Optima plate reader (excitation wavelength, 410 nm; 10-nm band-pass). The translocation rate was calculated as recommended in the LiveBLAzer FRET-B/G Loading Kit manual. Briefly, emission values were first corrected by subtraction of the average background signals recorded for empty wells, and then the 450-nm/520-nm emission ratio for each well was calculated. The translocation rate was expressed as the fold increase of the 450-nm/520-nm emission ratio of each infected cell in correlation with the emission ratio of uninfected cells. Experiments were performed three times.

*L. pneumophila* pXDC61 *ltpG* strains were grown for 16 h as described above but without the addition of IPTG. Expression of the TEM1 fusion of LtpG was subsequently induced for 5 h in bacterial liquid culture and, during the first hour of infection, by the addition of 1 mM IPTG.

To verify expression of the TEM1 fusion proteins in *Legionella*, 0.5 ml of the 21 h-cultures used for infection was harvested, and equal amounts of bacteria, adjusted according to the $OD_{600}$, were analyzed by Western blotting using a mouse anti-β-lactamase antibody (QED Bioscience Inc.).

**Prevalence screen.** Genomic DNA for 54 clinical and environmental *L. pneumophila* isolates was obtained from the Respiratory and Systemic Infection Laboratory, Health Protection Agency Centre for Infection, London, United Kingdom (see Table S1a in the supplemental material). These isolates were obtained from a range of locations across the United Kingdom and were previously characterized by serogroup, monoclonal antibody (MAb) subgroup, allelic profile, and sequence type (42, 43, 77). Thirty-three isolates were obtained from the Microbiological Diagnostic Unit at the University of Melbourne, as described previously (71), and from the collection of Stacey Yong, Taylor's University College, Malaysia (see Table S1b in the supplemental material). The PCR primers were designed to amplify 400- to 700-bp sequences of the putative effector genes and controls (Table 2). PCR screening using standard conditions was performed three times, and the results were analyzed by agarose gel electrophoresis. PCR-negative strains from the Australian and Malaysian collections were, in addition, confirmed by Southern hybridization, as described previously (71).

**Nucleotide sequence accession number.** The annotated draft genome sequence of *L. pneumophila* 130b can be obtained from the European Nucleotide Archive under accession number FR687201, and the raw sequence data can be obtained from the Sequence Read Archive under accession number ERA011231.

## RESULTS AND DISCUSSION

**General features of the *L. pneumophila* 130b genome.** The genome of *L. pneumophila* strain 130b was sequenced using 454 technology to approximately 11-fold coverage and assembled into 4 scaffolds (consisting of 145 contigs) and 14 small, unscaffolded contigs (Fig. 1). Analysis of the sequence suggested that the genome consists of a single circular chromosome of approximately 3.5 Mb with an average G+C content of 38%. No plasmids were identified in the genome sequence. The draft sequence was predicted to contain 3,293 coding sequences (CDSs) and 42 tRNAs, which is in good correlation

TABLE 2. Sequences of the primers used for the screen of *L. pneumophila* isolates

| Gene | Primer name | Primer sequence (5′–3′) |
|---|---|---|
| *dotA* (*lpw_29401*) | GNS068 DotA FW | ATTAGCTATTACGGTCCTCCTTTG |
| | GNS069 DotA RV | GAGTAGGATTACCCCCACAAG |
| *lpw_00581* (*ltpA*) | GNS078 seq FW | TCCTTAAAAAATTCTGATGTGCCTC |
| | GNS079 seq RV | CTTGTCTATATTGCTCGGGAGTAA |
| *lpw_02301* (*ltpB*) | GNS076 seq FW | ATGGCATCGGGGAAACAAGATA |
| | GNS077 seq RV | CCTCCTCCATCGTATCTTCAAAA |
| *lpw _02381* (*ltpC*) | PrimerRasGEF _F | ACGGTTCCCTTGGGTGAATTTC |
| | PrimerRasGEF _R | CAGCATATCTTCATCAGAAGGCTGG |
| *lpw _03701* (*ltpD*) | GNS109 seq FW | TTTTTCTTTCCTTCGCCGCAAG |
| | GNS110 seq RV | CCGTAAAGCCTCATGGCGTTC |
| *lpw _04551* (*ltpE*) | GNS063 seq FW | TCCATTTAATGGATTACCCAAAGAC |
| | GNS064 seq RV | GCAGCAAAAGCAGAACCTAATTTAA |
| *lpw _16311* (*ltpF*) | GNS074 seq FW | AAATGAGCTGATTAAGCGGATTAAC |
| | GNS075 seq RV | TTAGAATTTTACCCAGGCCATCA |
| *lpw _20091* (*ltpG*) | GNS070 seq FW | TGCTTTAGTGTATGGATTTGAACCA |
| | GNS071 seq RV | GGTTGAGGTTATATGCACTCGC |
| *lpw _20341* (*ltpH*) | GNS111 seq FW | CAAATGTTAAGAGCAGCAGCTG |
| | GNS112 seq RV | TCATAAAATAGAGCAGTAGGACCA |
| *lpw_25791* (*ltpI*) | GNS086 FW | CGCACTGAGGTACCCCTTTTCAACTTCCTTTTGAAAAA |
| | GNS099 RV | GACGTCGCATCTAGATTACATGAATCTTACAGAATGGCT |
| *lpw_26201* (*ltpJ*) | GNS072 seq FW | CTGAATTAGCTCAAGCTTTT |
| | GNS073 seq RV | TGAATTACCGATTCTTTGAGC |

with the other sequenced *L. pneumophila* strains (Philadelphia, 3.4 Mb, 2,942 CDSs, 43 tRNAs; Lens, 3.3 Mb, 2,947 CDSs, 43 tRNAs; Paris, 3.5 Mb, 3,082 CDSs, 43 tRNAs; Corby: 3.6 Mb, 3,204 CDSs, 44 tRNAs). The genome of 130b is highly syntenic with those of the other sequenced *L. pneumophila* strains, although, as can be expected from a species showing evidence of high genome plasticity (16), there are a number of insertions and deletions and other regions of difference (RODs) that are evident in whole-genome comparisons (data not shown).

**Relationship of 130b to other legionellae.** Phylogenetic analysis was performed using 26 genes from the five sequenced *L. pneumophila* serogroup 1 genomes (130b, Lens, Philadelphia-1, Corby, and Paris) and *L. longbeachae* NSW150. These phylogenetic markers—*frr*, *infC*, *nusA*, *pgk*, *rplA*, *rplB*, *rplC*, *rplD*, *rplE*, *rplK*, *rplL*, *rplM*, *rplN*, *rplP*, *rplS*, *rplT*, *rpmA*, *rpoB*, *rpsB*, *rpsC*, *rpsE*, *rpsI*, *rpsK*, *rpsM*, *smpB*, and *tsf*—are housekeeping genes that are involved in replication, transcription, translation, or central metabolism processes. These genes were used in a previous study to infer evolutionary relationships between 578 bacterial strains (96). These single-copy housekeeping genes were used to enable a phylogenetic signal to be detected from truly orthologous sequences free from horizontal gene transfer (HGT) or duplication events.

The resulting phylogenetic tree (Fig. 2) showed all the *L. pneumophila* serogroup 1 strains to be monophyletic, with *L. longbeachae* as an outgroup. Within the serogroup 1 clade, strains Lens and 130b were grouped together, with a very high bootstrap value (99.8%) in comparison with all the other *L. pneumophila* strains. This suggests that Lens and 130b have only recently diverged relative to the other *L. pneumophila* strains. This high-confidence grouping is an interesting result, as Lens was only recently isolated from an epidemic outbreak in France in 2003 and 2004, while 130b is a commonly used laboratory strain that originated from a transtracheal aspirate isolated in the United States in 1978 (16, 29, 30).

**The secretion systems of *L. pneumophila* 130b.** The main aim of sequencing the 130b genome was to facilitate the analysis of the secretion systems, in particular, the type IV secretion systems, and the Dot/Icm T4SS effector repertoire.

An overview of the protein secretion systems encoded by *L. pneumophila* 130b is shown in Table 3. The genome encodes the structural components of a putative type 1 secretion system (T1SS), which was first described as the *Legionella* secretion system (Lss) in *L. pneumophila* Corby (49), including a putative T1SS-associated regulatory GGDEF family protein, LssE. The functionality of the T1SS in *L. pneumophila* has yet to be demonstrated.

Functional analysis of the T2SS of *L. pneumophila* has been performed mainly in strain 130b; several substrates have been identified, and a contribution to *Legionella* virulence has been established (20, 79). The genes comprising the T2SS system are not in a single locus but form several clusters scattered around the genome, and this organization is largely conserved in location and sequence among the different *L. pneumophila* genomes, suggesting strong selection for its retention.

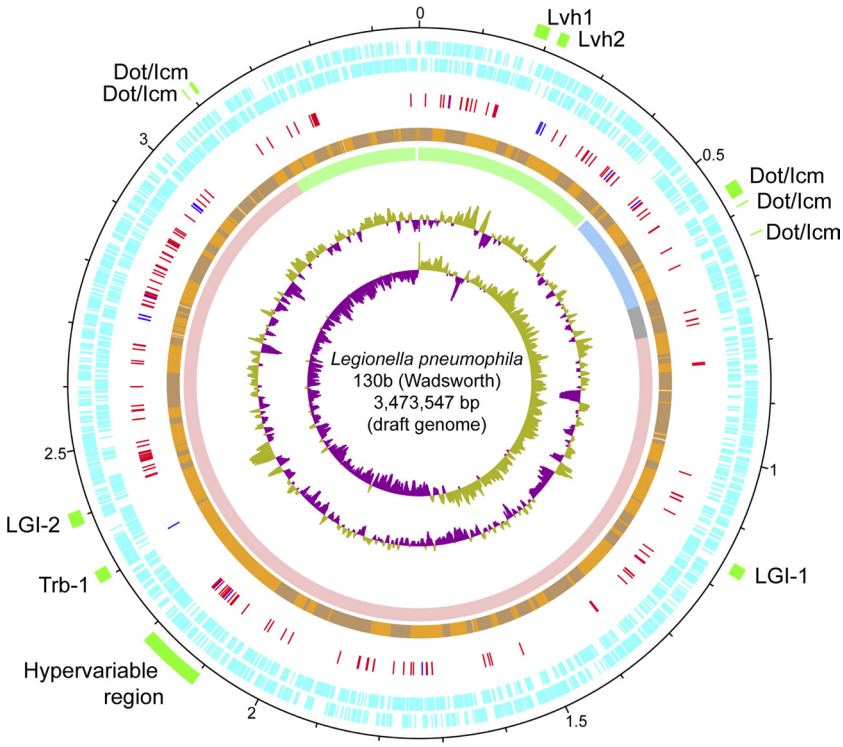


FIG. 1. Circular map of the *L. pneumophila* 130b draft genome. From the outside in, the green regions in the first circle show the positions of the T4SSs (detailed in Table 3) and the hypervariable region (shown in detail in Fig. 5). The second circle shows the scale in Mbp. The third and fourth circles show the predicted CDSs transcribed clockwise and counterclockwise, respectively. The fifth circle shows known *Legionella* T4SS effector genes (described in Table 4) and their paralogues (colored red) and new putative effectors (dark blue) (Table 5). The sixth circle shows the 159 contigs of the draft genome, colored brown and orange alternately, and circle 7 shows the 4 scaffolds that link 145 of these contigs (scaffold 1, light green; scaffold 2, light blue; scaffold 3, light gray; scaffold 4, light pink). Circle 8 shows a plot of the GC content (percent), and the innermost circle is a plot of GC deviation (G − C/G + C).

**Type IV secretion systems.** *L. pneumophila* 130b has the complete Dot/Icm T4SSB encoded in five loci in two distinct regions on the chromosome (Fig. 1). Although some degree of sequence variation in the Dot/Icm systems of different strains has been identified, the loci are largely conserved (Fig. 3A) (69, 86, 91) and are found in syntenic regions in each of the other sequenced *L. pneumophila* genomes.

Lvh T4SSA regions have been identified in the sequenced *L. pneumophila* strains Philadelphia-1, Paris, and Lens. The genes encoding the Lvh T4SSA are found either on a plasmid-like element or integrated in the chromosome (16, 18, 27). Furthermore, an *lvh* gene cluster was also found as part of a chromosomally integrated putative plasmid-like element in *L. longbeachae* strain D-4968, showing interspecies mobility of the *lvh* region (56). However, in strain 130b, we identified two distinct *lvh* regions, both of which were integrated into the chromosome (Fig. 1). We found no evidence in the sequence reads that either of these regions was mobilized on plasmid-like elements. Lvh1 (Fig. 3B) is almost identical (99% DNA identity), except for a few SNPs, to the published *lvh* region previously sequenced from strain 130b (accession number AF410854 [83]). The SNPs could be due to sequencing errors in either or both sequenced regions or they could be indicative of sequence divergence. The second T4SSA Lvh system in 130b, Lvh2, has 99% sequence identity to the Lvh region in the genome of *L. pneumophila* Paris. Lvh1 and Lvh2 in 130b have only 92% similarity to each other, and therefore, it seems unlikely that the two Lvh regions arose from a duplication event. Instead, it seems more likely that Lvh2 was acquired via a plasmid-like element highly similar to that found in the Paris strain and was subsequently integrated into the chromosome. The two Lvh loci are located close to each other in the 130b genome (Fig. 1), on either side of the CRISPR locus, and are divergently transcribed. Both are part of a larger ROD that shows variation in each of the sequenced strains.

The Lvh T4SS is not present in *L. pneumophila* strain Corby. Instead, two similar T4SSAs, Trb-1 and Trb-2, are encoded on separate genomic islands in the strain (37). Both genomic islands can exist integrated into the genome or as episomal plasmids. Glöckner et al. and Samrakandi et al. independently suggested that *L. pneumophila* 130b contains a Trb-1-like T4SSA (37, 83). The 130b genome sequence confirmed the presence of a T4SSA gene cluster (*lpw_22591* to *lpw_22831*) that shows similarity to the *L. pneumophila* Corby Trb-1 T4SSA (Fig. 3C); however, the genetic context in 130b differs (37). Integration of the *L. pneumophila* Corby Trb-1 genomic island was reported to be specifically into the $tRNA^{Pro}$ gene. In contrast, in *L. pneumophila* 130b, the 22-kb Trb-1-like cluster is found in a nonsynonymous genomic location as part of a larger, approximately 150-kb ROD, and there was no evidence in the sequence reads that it can exist in plasmid-like form in 130b.

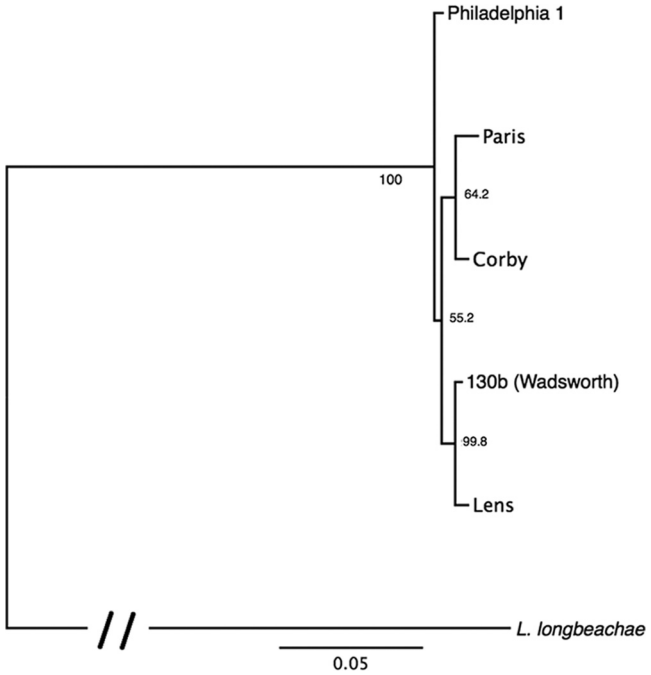


FIG. 2. Phylogeny of *Legionella* showing the phylogenetic relationship of *L. pneumophila* 130b to the other sequenced *L. pneumophila* strains, Corby, Paris, Lens, and Philadelphia-1, with *L. longbeachae* NSW150 as an outgroup. The tree was built using 26 housekeeping genes (*frr*, *infC*, *nusA*, *pgk*, *rplA*, *rplB*, *rplC*, *rplD*, *rplE*, *rplK*, *rplL*, *rplM*, *rplN*, *rplP*, *rplS*, *rplT*, *rpmA*, *rpoB*, *rpsB*, *rpsC*, *rpsE*, *rpsI*, *rpsK*, *rpsM*, *smpB*, and *tsf*) that were found in single copies in all six genomes. The numbers next to the branches show percentages of support from a bootstrap analysis of 500 replicates. The tree shows a highly supported (99.8%) grouping of Lens and 130b. The scale bar represents the number of substitutions per site.

In addition to the Dot/Icm T4SSB and the three T4SSA systems, we found two highly similar clusters of 24 genes in the *L. pneumophila* 130b genome, which resemble clusters of the recently defined GI-T4SS (50). We named these putative new T4SSs *Legionella* GI-T4SS-1 (LGI-1; *lpw_10731* to *lpw_10961*) and *Legionella* GI-T4SS-2 (LGI-2; *lpw_21631* to *lpw_21861*) (Fig. 1). Both the LGI-1 and LGI-2 clusters encode homologues of the T4SS components VirD4, VirB4, and TraB and several proteins that have significant similarity to the proteins encoded in the transfer region of several integrative and conjugative elements (ICEs) (Fig. 4), including ICEhin1056 from *Haemophilus influenzae* (accession number AJ627386 [67]). The first four genes of both LGI clusters encode homologues of the proteins LvrR, LvrA, LvrB, and the CsrA-like global regulator protein LvrC, all of which are found in analogous positions in the Lvh and the Trb-1-like T4SSA systems. This suggests an important role for these proteins and a common mechanism for the regulation of T4SSA and GI-T4SS in *L. pneumophila*.

Although it was not previously recognized as a putative T4SS cluster, we found that LGI-1 was conserved in the other sequenced *L. pneumophila* strains and was in the same genomic location, although there is variation in the gene content immediately flanking LGI-1 in Philadelphia-1, Paris, and Corby. This region was previously found to be heterogeneously distributed among 217 *L. pneumophila* strains (15, 18), and based on this observation, it was suggested that the region might be mobile.

LGI-2 is part of a 150-kb ROD in 130b, which also encodes the Trb-1-like T4SSA, which is absent from strains Lens and Philadelphia. LGI-2 is highly similar (approximately 87% DNA identity) to regions in the genomes of strains Corby (*lpc1857* to *lpc1880*) and Paris (*lpp2375* to *lpp2398*). However, these LGI-2-like regions are found in a different genomic lo-

TABLE 3. Overview of the *L. pneumophila* strain 130b secretion systems

| Secretion system | Gene | Locus |
|---|---|---|
| General secretory (Sec) system | *secA*, *-B* | *lpw_14821*, *lpw_24951* |
| | *secE*, *-Y* | *lpw_03951*, *lpw_04281* |
| | *yajC*, *secD*, *-F* | *lpw_20591* to *lpw_20611* |
| | *secG*, *yidC* | *lpw_30451*, *lpw_32881* |
| Twin arginine translocation (TAT) system | *tatA*, *-B*, *-C* | *lpw_31791*, *lpw_31801*, *lpw_32171* |
| T1SS | *lssXYZABDE* | *lpw_15351* to *lpw_15411* |
| T2SS | *pilD*, *lspK* to *-F* | *lpw_15471*, *lpw_13701* to *lpw_13651* |
| | *lspE*, *-D*, *lspC* | *lpw_13281*, *lpw_13291*, *lpw_09771* |
| | *lepB*, *lspL*, *lspM* | *lpw_19051*, *lpw_19081*, *lpw_19091* |
| T4SS | | |
| Dot/Icm T4SSB | *dotA*, *icmV*, *-W*, *-X* | *lpw_29401* to *lpw_29431* |
| | *dotD*, *-C*, *-B* | *lpw_29281* to *lpw_29301* |
| | *icmT*, *-S*, *-R*, *-Q*, *-P*, *-O*, *-N*, *-M*, *-L*, *-K*, *-E*, *-G*, *-C*, *-D*, *-J*, *-B*, *-F*, *-H* | *lpw_05221* to *lpw_05401* |
| | *dotV*, *lvgA* | *lpw_05521*, *lpw_06041* |
| Lvh1 T4SSA | *lvr1E* to *lvr1A* | *lpw_01541* to *lpw_01691* |
| Lvh2 T4SSA | *lvr2A* to *lvr2E* | *lpw_01881* to *lpw_02031* |
| Trb-1-like T4SSA | *lvhrI-traM1* | *lpw_22591* to *lpw_22831* |
| GI-T4SS 1 | *Lvr3R*, *-A* to *-C*, *lgi1A* to *-T* | *lpw_10731* to *lpw_10961* |
| GI-T4SS 2 | *Lvr4R*, *-A* to *-C*, *lgi2A* to *-T* | *lpw_21631* to *lpw_21861* |

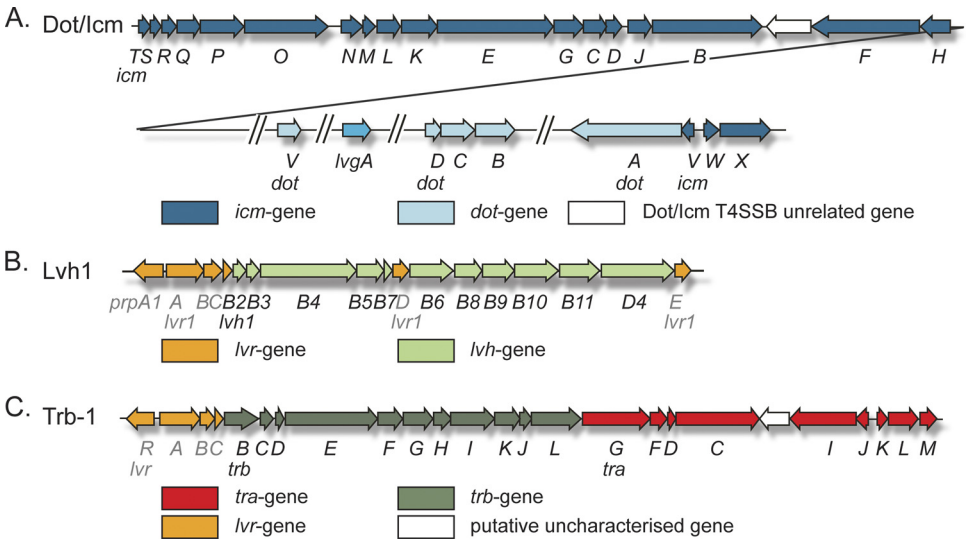


FIG. 3. Organization of the Dot/Icm T4SSB (A), Lvh1 T4SSA (B), and Trb1-like T4SSA (C) loci on the *L. pneumophila* 130b chromosome. The color coding represents shared gene names.

cation, and the flanking regions are highly divergent in all three strains.

Even though the functionality of the two putative new T4SSs has yet to be demonstrated, it is tempting to speculate that they might contribute to the genetic competence and mobilization of genomic islands in *L. pneumophila*, adding another layer of complexity to the factors that shape the genome plasticity and virulence of *L. pneumophila*.

**The conserved Dot/Icm T4SS effectors.** *L. pneumophila* has accumulated a large number of T4SS effector proteins. To date, Dot/Icm-dependent translocation has been experimentally proven for 151 effector proteins, and several dozen additional candidate Dot/Icm effectors have been identified by sequence analysis or specialized assays (9, 24, 33, 57, 63, 87, 98). Furthermore, it has been estimated that the total number of effector proteins could be as high as 300 (9). However, only a fraction of the proven effector proteins have been functionally characterized so far, and this analysis is hampered by families of effector protein paralogues that have presumed functional redundancy.

Homologues of 136 of the 151 proven effectors were identified in the *L. pneumophila* 130b genome. All of the effectors that are not present in strain 130b are also absent from at least one of the other sequenced *L. pneumophila* strains (Table 4). It is notable that no effector pseudogenes were identified in the 130b genome, whereas at least one effector pseudogene was identified in each of the other *L. pneumophila* genomes, which totaled 15 effector pseudogenes across all four strains compared (Table 4). A core set of 107 intact effector genes is common to all five sequenced *L. pneumophila* genomes. The conservation of this core set of T4SS effectors could be indicative of their biological importance. Recently, the genome sequences of the *L. longbeachae* strains D-4968 and NSW150 were reported (14, 56). *L. longbeachae* encodes a Dot/Icm T4SSB and replicates in a specialized vacuole in macrophages, which shares features with, but is distinct from, the *L. pneumophila* replicative vacuole (4). *L. pneumophila* and *L. longbeachae* encode a common set of effectors, which includes about one-third of the validated *L. pneumophila* effectors and may constitute the minimum set of effectors required by *Legionella* species to survive and replicate in macrophages. Once more *L. longbeachae* strains and other *Legionella* species, such as *Legionella micdadei*, are sequenced, the core set of essential effectors will be defined with greater confidence.

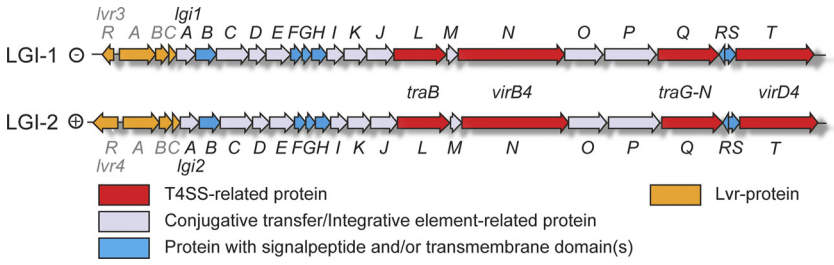


FIG. 4. Organization of the putative genomic-island-associated T4SS gene clusters LGI-1 and LGI-2 on the *L. pneumophila* 130b chromosome. The color coding shows the sequence and predicted structural homology of the encoded proteins to T4SS-associated proteins, conjugative-transfer/integrative-element-associated proteins, and the *Legionella* Lvr proteins or the presence of transmembrane domains and/or signal peptides that are conserved between the two clusters (blue). The scheme illustrates the high degree of organizational conservation of the two clusters. Plus and minus indicate that the clusters are on opposite strands of the chromosome.

TABLE 4. Presence of known *L. pneumophila* T4SS effectors in sequenced genomes

| No. | Name[a] | Synonym | Gene in[b]: | | | | |
|---|---|---|---|---|---|---|---|
| | | | 130b | Lens | Philadelphia | Corby | Paris |
| 1 | CegC1 | | *lpw_00111* | *lpl0012* | *lpg0012* | *lpc0013* | *lpp0012* |
| 2 | LegA10 | AnkQ | *lpw_00381* | *lpl0038* | *lpg0038* | *lpc0039* | *lpp0037* |
| 3 | lpg0045 | | *lpw_00441* | *lpl0044* | *lpg0045* | *lpc0047* | *lpp0046* |
| 4 | Ceg3 | | *lpw_00781* | | *lpg0080* | | *lpp0094* |
| 5 | lpg0081 | | *lpw_00791* | | *lpg0081* | | *lpp0095* |
| 6 | Lem1 | | *lpw_00881* | *lpl0089* | *lpg0090* | *lpc0109* | *lpp0104* |
| 7 | Ceg4 | | *lpw_00961* | *lpl0096* | *lpg0096* | *lpc0115* | *lpp0110* |
| 8 | VipF | | *lpw_01031* | *lpl0103* | *lpg0103* | *lpc0122* | *lpp0117* |
| 9 | CegC2 | | *lpw_01261* | *lpl0125* | *lpg0126* | *lpc0146* | *lpp0140* |
| 10 | SdhB | | *lpw_01361* | *lpl0135* | *lpg0135* | *lpc0156* | *lpp0150* |
| 11 | LegU1 | | *lpw_02651* | *lpl0234* | *lpg0171* | *lpc0252* (*p*) | *lpp0233* |
| 12 | Ceg5 | | *lpw_02821* | | *lpg0191* | | *lpp0251* |
| 13 | Ceg7 | | *lpw_03151* | *lpl0281* | *lpg0227* | *lpc0303* | *lpp0286* |
| 14 | SidE | | *lpw_03221* | *lpl0288* | *lpg0234* | *lpc0309* | *lpp0304* |
| 15 | Ceg8 | | *lpw_03291* | *lpl0294* | *lpg0240* | *lpc0316* | *lpp0310* |
| 16 | Ceg9 | | *lpw_03361* | *lpl0300* | *lpg0246* | *lpc0323* | *lpp0316* |
| 17 | SdbA | | *lpw_03641* | *lpl0327* | *lpg0275* | *lpc0351/2* (*p*) | *lpp0349* |
| 18 | LegG2 | | *lpw_03651* | *lpl0328* | *lpg0276* | *lpc0353* | *lpp0350* |
| 19 | Ceg10 | | *lpw_03741* | *lpl0336* | *lpg0284* | *lpc0361* | *lpp0360* |
| 20 | Lem2 | | *lpw_03751* | *lpl0337* | *lpg0285* | *lpc0362* | *lpp0361* |
| 21 | lpg0294 | | *lpw_03861* | *lpl0347* | *lpg0294* | *lpc0373* | *lpp0372* |
| 22 | lpg0365 | | *lpw_04441* | *lpl0406* | *lpg0365* | *lpc2979* | *lpp0430* |
| 23 | SdhA | | *lpw_04591* | *lpl0419* | *lpg0376* | *lpc2967* | *lpp0443* |
| 24 | VipA | | *lpw_04721* | *lpl0433* | *lpg0390* | *lpc2954* | *lpp0457* (*p*) |
| 25 | LegA9 | AnkY | | | *lpg0402* | | *lpp2058* |
| 26 | LegA7 | AnkG, AnkZ | *lpw_04841* | *lpl0445* | *lpg0403* | *lpc2941* | *lpp0469* |
| 27 | LegA11 | AnkJ | *lpw_05181* | *lpl0479* | *lpg0436* | *lpc2906* | *lpp0503* |
| 28 | Ceg14 | | *lpw_05191* | *lpl0480* | *lpg0437* | *lpc2905* | *lpp0504* |
| 29 | LegA12 | AnkC | *lpw_05631* | *lpl0523* | *lpg0483* | *lpc2861* | *lpp0547* |
| 30 | lpg0518 | | *lpw_05981* | *lpl0557* | *lpg0518* | *lpc2826* | *lpp0581* |
| 31 | Ceg17 | | | | *lpg0519* | | |
| 32 | SidA | | *lpw_06951* | *lpl0658* | *lpg0621* | *lpc2673* | *lpp0675* |
| 33 | lpg0634 | | *lpw_07081* | *lpl0671* | *lpg0634* | *lpc2660* | *lpp0688* |
| 34 | WipB | | *lpw_07161* | *lpl0679* | *lpg0642* | *lpc2651* | *lpp0696/7* (*p*) |
| 35 | LegA8 | AnkN, AnkX | *lpw_07721* | *lpl0732* | *lpg0695* | *lpc2599* | *lpp0750* |
| 36 | Lem3 | | *lpw_07731* | *lpl0733* | *lpg0696* | *lpc2598* | *lpp0751* |
| 37 | Ceg18 | | *lpw_09801* | *lpl0929* | *lpg0898* | *lpc2395* | *lpp0959* |
| 38 | LidA | | *lpw_10251* | *lpl0971* | *lpg0940* | *lpc2349* | *lpp1002* |
| 39 | LegL1 | | *lpw_10311* | | *lpg0945* | *lpc2344* | *lpp1007* |
| 40 | lpg0963 | LigA | *lpw_10491* | *lpl0992* | *lpg0963* | *lpc2324* | *lpp1025* |
| 41 | Lem4 | | *lpw_11451* | *lpl1100* | *lpg1101* | *lpc2154* (*p*) | *lpp1101* |
| 42 | Lem5 | | *lpw_11571* | *lpl1114* | *lpg1110* | *lpc2142* | *lpp1111* |
| 43 | Lem6 | | *lpw_11681* | | *lpg1120* | | |
| 44 | Ceg19 | | *lpw_11691* | *lpl1126* | *lpg1121* | *lpc0578* | *lpp1121* |
| 45 | CegC3 | | *lpw_11971* | *lpl1150* | *lpg1144* | *lpc0607* | *lpp1146* |
| 46 | Lem7 | | *lpw_11981* | *lpl1151* | *lpg1145* | *lpc0608* | *lpp1147* |
| 47 | lpg1148 | | *lpw_12011* | *lpl1154* | *lpg1148* | *lpc0611* | *lpp1150* |
| 48 | lpg1158 | | *lpw_12121* | *lpl1165* (*p*) | *lpg1158* | *lpc0621* | *lpp1160* |
| 49 | VpdB | | *lpw_12861* | *lpl1235* | *lpg1227* | *lpc0696* | *lpp1235* |
| 50 | lpg1273 | | *lpw_12871* | *lpl1236* | *lpg1273* | *lpc0698* | *lpp1236* |
| 51 | Lem8 | | | | *lpg1290* | | *lpp1253* |
| 52 | SidG | | | | *lpg1355* | | *lpp1309* |
| 53 | VpdC | | *lpw_14431* | *lpl1377* | *lpg1426* | *lpc0842* | *lpp1381* |
| 54 | LegK1 | | *lpw_15031* | *lpl1545* | *lpg1483* | *lpc0898* | *lpp1439* |
| 55 | LegC5 | Lgt3 | *lpw_15081* | *lpl1540* | *lpg1488* | *lpc0903* (*p*) | *lpp1444* |
| 56 | Lem9 | | | | *lpg1491* | | *lpp1447* |
| 57 | Lem10 | | *lpw_15181* | *lpl1530* | *lpg1496* | *lpc0915* | *lpp1453* |
| 58 | LegC6 | | *lpw_16131* | *lpl1437* | *lpg1588* | *lpc1013* | *lpp1546* |
| 59 | Lem11 | | *lpw_16231* | *lpl1427* | *lpg1598* | *lpc1025* | *lpp1556* |
| 60 | LegL2 | | *lpw_16241* | *lpl1423* (*p*) | *lpg1602* | *lpc1028* | *lpp1567* |

*Continued on following page*

TABLE 4—*Continued*

| No. | Name[a] | Synonym | Gene in[b]: | | | | |
|---|---|---|---|---|---|---|---|
| | | | 130b | Lens | Philadelphia | Corby | Paris |
| 61 | Ceg23 | | *lpw_16461* | *lpl1402* | *lpg1621* | *lpc1048* | *lpp1591* |
| 62 | Lem12 | | *lpw_16511* | *lpl1398* | *lpg1625* | *lpc1052* | *lpp1595* |
| 63 | SidB | | *lpw_16681* | *lpl1384* | *lpg1642* | *lpc1071* | *lpp1612* (*p*) |
| 64 | LegL3 | | *lpw_16861* | *lpl1625* | *lpg1660* | *lpc1090* | *lpp1631* |
| 65 | lpg1689 | | *lpw_17141* | *lpl1652* | *lpg1689* | *lpc1120* | *lpp1658* |
| 66 | LegC3 | PpeA | *lpw_17231* | *lpl1660* | *lpg1701* | *lpc1130* | *lpp1666* |
| 67 | PpeB | Lem13 | *lpw_17241* | *lpl1661* | *lpg1702* | *lpc1131* | *lpp1667* |
| 68 | lpg1717 | | *lpw_17401* | | *lpg1717* | | *lpp1682* |
| 69 | Lpl1680 | | | *lpl1680* | | *lpc1150* | |
| 70 | Lpl1681 | | | *lpl1681* | | *lpc1151* | |
| 71 | LegAS4 | AnkI | *lpw_17411* | *lpl1682* | *lpg1718* | *lpc1152* | *lpp1683* |
| 72 | Lpg1751 | | *lpw_17761* | *lpl1715* | *lpg1751* | *lpc1191* | *lpp1715* |
| 73 | Ceg25 | | *lpw_18691* | *lpl1800* | *lpg1836* | *lpc1280* | *lpp1799* |
| 74 | Lem14 | | *lpw_18871* | *lpl1817* | *lpg1851* | *lpc1296* | *lpp1818* |
| 75 | YlfB | LegC2 | *lpw_19161* | *lpl1845* | *lpg1884* | *lpc1331* | *lpp1848* |
| 76 | LegLC8 | | *lpw_19231* | *lpl1852* | *lpg1890* | *lpc1338* | *lpp1857* |
| 77 | Lem15 | | *lpw_19721* | *lpl1903* | *lpg1933* | *lpc1406* | *lpp1914* |
| 78 | Lem16 | | *lpw_19951* | *lpl1917b* (*p*) | *lpg1947* | *lpc1421b* (*p*) | *lpp1930* |
| 79 | LegLC4 | | | | *lpg1948* | | |
| 80 | Lem17 | | *lpw_19961* | *lpl1918* | *lpg1949* | *lpc1422* | *lpp1931* |
| 81 | RalF | | *lpw_19971* | *lpl1919* | *lpg1950* | *lpc1423* | *lpp1932* |
| 82 | LegC4 | | *lpw_20041* | *lpl1922* | *lpg1953* | *lpc1426* | *lpp1935* |
| 83 | LegL5 | | | | *lpg1958* | | *lpp1940* |
| 84 | LirA | | *lpw_20111* | *lpl1934* (*p*) | *lpg1960* | *lpc1437* | *lpp1942* |
| 85 | LirB | | *lpw_20131* | *lpl1936* | *lpg1962* | *lpc1440* | *lpp1946* |
| 86 | LirC | PieA | | | *lpg1963* | *lpc1442* | |
| 87 | LirD | PieB | | | *lpg1964* | | |
| 88 | LirE | PieC | *lpw_20141* | | *lpg1965* | *lpc1443/5* (*p*) | |
| 89 | LirF | PieD | *lpw_20151* | | *lpg1966* | *lpc1446* | *lpp1947* |
| 90 | PieE | Lem18 | *lpw_20201* | *lpl1941* | *lpg1969* | *lpc1452* | *lpp1952* |
| 91 | PieF | | *lpw_20291* | *lpl1950* | *lpg1972* | *lpc1459* | *lpp1955* |
| 92 | LegG1 | PieG | *lpw_20351* | *lpl1953* | *lpg1975/6* (*p*) | *lpc1462* | *lpp1959* |
| 93 | SetA | | *lpw_20371* | *lpl1955* | *lpg1978* | *lpc1464* | *lpp1961* |
| 94 | LegK2 | | *lpw_23101* | *lpl2066* | *lpg2137* | *lpc1586* | *lpp2076* |
| 95 | LegAU13 | Ceg27, AnkB | *lpw_23181* | *lpl2072* | *lpg2144* | *lpc1593* | *lpp2082* |
| 96 | SdeC | | *lpw_23271* | *lpl2081* | *lpg2153* | *lpc1602* | *lpp2092* |
| 97 | Sde | | *lpw_23281* | *lpl2082* | *lpg2154* | *lpc1603* | *lpp2093* |
| 98 | SidJ | | *lpw_23291* | *lpl2083* | *lpg2155* | *lpc1604* | *lpp2094* |
| 99 | SdeA | | *lpw_23331* | *lpl2085* | *lpg2157* | *lpc1618* | *lpp2096* |
| 100 | Lpg2160 | | *lpw_23361* | *lpl2088* | *lpg2160* | *lpc1621* | *lpp2099* |
| 101 | Lem19 | | *lpw_23451* | *lpl2093* | *lpg2166* | *lpc1626* | *lpp2104* |
| 102 | LegS2 | | *lpw_23561* | *lpl2102* | *lpg2176* | *lpc1635* | *lpp2128* |
| 103 | CegC4 | | *lpw_23821* | *lpl2124* | *lpg2200* | *lpc1664* | *lpp2150* |
| 104 | LegA2 | | *lpw_24011* | *lpl2140* | *lpg2215* | *lpc1680* | *lpp2166* |
| 105 | Lem20 | | *lpw_24021* | *lpl2141* | *lpg2216* | *lpc1681* | *lpp2167* |
| 106 | LpnE | | *lpw_24081* | *lpl2147* | *lpg2222* | *lpc1689* | *lpp2174* |
| 107 | PpgA | | | | *lpg2224* | | |
| 108 | Lem21 | | *lpw_24371* | *lpl2174* | *lpg2248* | *lpc1717* | *lpp2202* |
| 109 | YlfA | LegC7 | *lpw_24841* | *lpl2217* | *lpg2298* | *lpc1763* | *lpp2246* |
| 110 | LegA3 | AnkH, | *lpw_24871* | *lpl2219* | *lpg2300* | *lpc1765* | *lpp2248* |
| 111 | LegA5 | AnkK | *lpw_25121* | *lpl2242* | *lpg2322* | *lpc1789* | *lpp2270* |
| 112 | Lpg2327 | | *lpw_25181* | *lpl2247* | *lpg2327* | *lpc1794* | *lpp2275* |
| 113 | Lem22 | | *lpw_25191* | *lpl2248* | *lpg2328* | *lpc1795* | *lpp2276* |
| 114 | SdbC | | *lpw_26021* | *lpl2315* | *lpg2391* | *lpc2086* | *lpp2458* |
| 115 | LegL7 | | *lpw_26121* | *lpl2323* | *lpg2400* | | |
| 116 | Lem23 | | *lpw_26191* | *lpl2329* | *lpg2406* | *lpc2070* | *lpp2472* |
| 117 | Lpg2407 | | | | *lpg2407* | *lpc2069* | *lpp2474* |
| 118 | Ceg29 | | *lpw_26241* | *lpl2332* | *lpg2409* | *lpc2067* | *lpp2476* |
| 119 | VpdA | | *lpw_26261* | *lpl2334* | *lpg2410* | *lpc2065* | *lpp2479* |
| 120 | Lem24 | | *lpw_26281* | *lpl2335* | *lpg2411* | *lpc2064* | *lpp2480* |

*Continued on following page*

TABLE 4—*Continued*

| No. | Name[a] | Synonym | Gene in[b]: | | | | |
|---|---|---|---|---|---|---|---|
| | | | 130b | Lens | Philadelphia | Corby | Paris |
| 121 | Lpp2486 | | | | | | *lpp2486* |
| 122 | Lem25 | | *lpw_26401* | *lpl2345* | *lpg2422* | *lpc2055* | *lpp2487* |
| 123 | Ceg30 | | *lpw_26521* | *lpl2353* | *lpg2433* | *lpc2043* | *lpp2500* |
| 124 | LegA14 | AnkF, Ceg31 | *lpw_26701* | *lpl2370* | *lpg2452* | *lpc2026* | *lpp2517* |
| 125 | LegA15 | AnkD | *lpw_26751* | *lpl2375* | *lpg2456* | *lpc2020* | *lpp2522* |
| 126 | SidM | DrrA | *lpw_26851* | *lpl2384* | *lpg2464* | | |
| 127 | SidD | | *lpw_26861* | *lpl2385* | *lpg2465* | | |
| 128 | SdbB | | *lpw_27041* | *lpl2402* | *lpg2482* | *lpc1996* | *lpp2546* |
| 129 | LepB | | *lpw_27131* | *lpl2411* | *lpg2490* | *lpc1987* | *lpp2555* |
| 130 | Ceg32 | SidI | *lpw_27301* | *lpl2426* | *lpg2504* | *lpc1967* | *lpp2572* |
| 131 | SdjA | | *lpw_27341* | *lpl2430* | *lpg2508* | *lpc1962/3* (*p*) | *lpp2576* |
| 132 | SdeD | | *lpw_27351* | *lpl2431* | *lpg2509* | *lpc1961* | *lpp2577* |
| 133 | SdcA | | *lpw_27361* | *lpl2432* | *lpg2510* | *lpc1960* | *lpp2578* |
| 134 | SidC | | *lpw_27371* | *lpl2433* | *lpg2511* | *lpc1959* | *lpp2579* |
| 135 | Lem26 | | *lpw_27501* | | *lpg2523* | | |
| 136 | lpg2527 | | *lpw_27531* | *lpl2447* | *lpg2527* | *lpc1944* | *lpp2592* |
| 137 | Lem27 | | *lpw_27551* | *lpl2449* | *lpg2529* | *lpc1942* | *lpp2594* |
| 138 | LegK3 | | *lpw_27911* | *lpl2481* | *lpg2556* | *lpc1906* | *lpp2626* |
| 139 | SidF | | *lpw_28321* | *lpl2507* | *lpg2584* | *lpc0561* | *lpp2637* |
| 140 | Ceg33 | | *lpw_28391* | *lpl2514* | *lpg2591* | *lpc0551* | *lpp2644* |
| 141 | Lem28 | | *lpw_28521* | *lpl2526* | *lpg2603* | *lpc0539* | *lpp2656* |
| 142 | WipA | | *lpw_29771* | *lpl2646* | *lpg2718* | *lpc0415* | *lpp2775* |
| 143 | Lpg2744 | | *lpw_30031* | *lpl2669* | *lpg2744* | *lpc0386* | *lpp2800* |
| 144 | LepA | | *lpw_30471* | *lpl2708* | *lpg2793* | *lpc3079* | *lpp2839* |
| 145 | Lem29 | | *lpw_30591* | *lpl2719* | *lpg2804* | *lpc3090* | *lpp2850* |
| 146 | Ceg34 | | *lpw_30831* | *lpl2714* | *lpg2826* | *lpc3113* | |
| 147 | SidH | | *lpw_30861* | | *lpg2829* | | |
| 148 | LegU2 | LubX | *lpw_30881* | | *lpg2830* | | *lpp2887* |
| 149 | VipD | | *lpw_30891* | | *lpg2831* | | *lpp2888* |
| 150 | LegC8 | Lgt2 | | | *lpg2862* | | |
| 151 | LegP | | *lpw_32851* | *lpl2927* | *lpg2999* | *lpc3315* | *lpp3071* |

[a] Where a gene name is not assigned, the locus tag from the genome in which the effector was initially identified is used.
[b] (p), partial or pseudogene.

**The effector-rich hypervariability region of *L. pneumophila*.** Approximately 30% of the proven effector proteins are absent from at least one of the sequenced *L. pneumophila* strains (Table 4). This substantial number of "accessory" effectors is one possible explanation for phenotypic differences in some virulence-associated attributes, such as host cell adherence, intracellular trafficking, and cytotoxicity, that have been reported between *L. pneumophila* 130b, *L. pneumophila* Philadelphia JR32, and other *L. pneumophila* strains (1, 83).

Two regions of high genomic plasticity, rich in effector proteins, were previously described in *L. pneumophila* (72, 98). These regions partially overlap (Fig. 5) and, in strain 130b, together comprise a 37-kb region located between *lpw_20071* and *lpw_20361* (Fig. 1). Moreover, whole-genome alignment revealed that this combined region forms part of a larger genomic region that displays considerable divergence among the five sequenced genomes (Fig. 5) and, in 130b, constitutes the 96 kb from *lpw_19681* to *lpw_20471*. In contrast, there are no gene deletions or insertions for more than 30 kb upstream and 48 kb downstream of the region shown in Fig. 5 (data not shown). This suggests that the two regions initially described by Zusman et al. and Ninio et al. (72, 98) constitute the inner core of a much larger 80- to 100-kb region of high genomic plasticity that represents a strain-specific variable effector region (Fig. 5). Adjacent to this large variable region is a conserved $tRNA^{Phe}$ gene, which could represent an insertion site if the region was horizontally acquired. However, no repeat ends or integrase genes were identified in any of the strains, and the presence of a distinct genomic island was not evident from examination of the G+C content of this region in each of the genomes. Notably, many of the insertions and deletions were found to be associated with insertion sequence (IS) elements, as highlighted for 130b in Fig. 5, and it is highly likely that these repetitive elements have contributed to the gene flux in this hypervariable region. Many of the putative CDSs in this hypervariable region encode hypothetical proteins, and therefore, it is possible that there are more effectors encoded in the region.

**Novel T4SS 130b effector proteins.** Prior to this work, 151 proteins were proven to be substrates of the Dot/Icm T4SS, a number that is unmatched by other bacterial pathogens. Most of the analysis was done using the Philadelphia strain. However, given the high genomic plasticity of the sequenced *L. pneumophila* strains, this is likely to result in an underestimation of the true number of effectors found in the species. Indeed, strain-specific effector candidates have been predicted in *L. pneumophila* Lens and Paris, but only three of them have been proven to be translocated (16, 33, 62). Accordingly, we

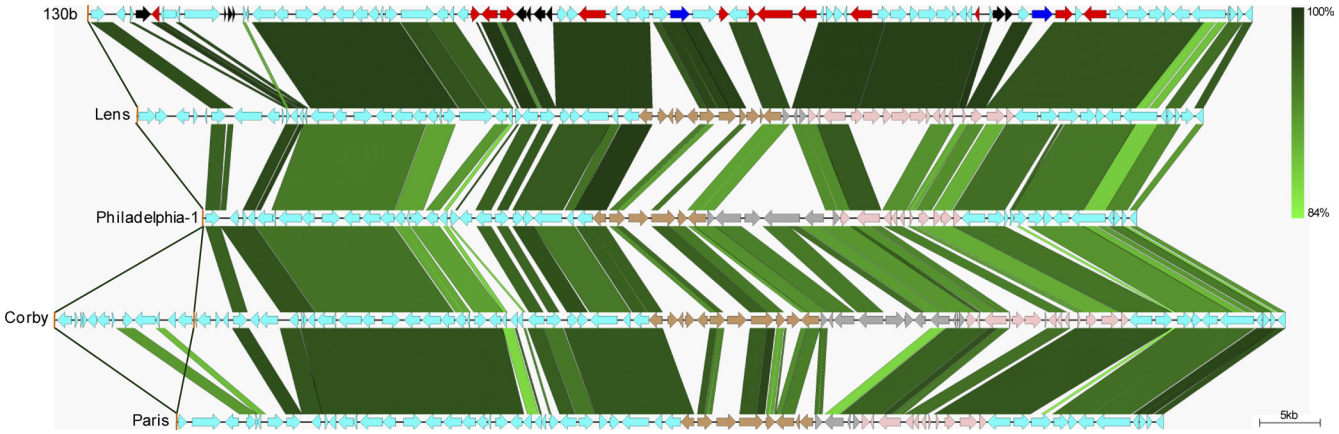


FIG. 5. A region of high genome variability in *L. pneumophila* encodes several new and known Dot/Icm T4SS effectors. Shown is a genome comparison of the five sequenced *L. pneumophila* genomes (130b, Lens, Philadelphia-1, Corby, and Paris). The CDSs for each genome are shown (130b *lpw_19681* to *lpw_20471*, Lens *lpl1893* to *lpl1966*, Philadelphia-1 *lpg1930* to *lpg1990*, Corby *lpc1384* to *lpc1473*, and Paris *lpp1904* to *lpp1971*). For 130b, CDSs encoding homologues of known effectors are colored red (from left to right, *lem15*, *lem16*, *lem17*, *ralF*, *legC4*, *lirAB*, *pieCDEFG*, and *setA*), new effectors are dark blue (*ltpG* and *ltpH*), and transposases are black. tRNA-Phe is represented by an orange bar. Regions with significant nucleotide similarity between the genomes are linked by shading (the percent identity [BLASTn] is indicated on the right). The scale bar indicates genome length. The genome comparison shows that the previously defined hypervariable effector region (98) (brown CDSs) and plasticity island of effectors (72) (pink CDSs) partially overlap (gray CDSs) and can be significantly extended.

analyzed the genome of *L. pneumophila* strain 130b for proteins with characteristics of known T4SS effectors, such as eukaryotic-like domains, that are absent from *L. pneumophila* Philadelphia-1. Our screen identified 16 new putative Dot/Icm T4SS effector proteins (Table 5).

Two of these new putative effector proteins, Lpw_20091 and Lpw_20341, are encoded in the inner core of the hypervariable effector-rich region (Fig. 5). The putative effector *lpw_20091* encodes a protein containing a predicted filamentation induced by the cyclic AMP (cAMP) (FIC) domain. Homologues of *lpw_20091* are also found in the same genomic location in Lens and Corby, but not in Philadelphia-1 or Paris (Fig. 5 and Table 5), which instead contain the effector gene *legL5* (*lpg1958* and *lpp1940*, respectively) at this location. The FIC domain was initially identified in the *Escherichia coli* cell division protein, Fic-1 (55). Structural similarity to FIC domain proteins was shown for the so-called death on curing (DOC) protein, which is part of a toxin-antidote plasmid addiction system in *E. coli* (54, 60). FIC domain proteins have recently gained increased attention, as members of this diverse family were found to be translocated virulence factors of several pathogenic bacteria (81, 94, 97). The FIC domain catalyzes the covalent modification of target proteins with AMP, interfering with host cell signaling. The conserved Dot/Icm T4SS effector LegA8/AnkX/AnkN (Lpg0695) belongs to the FIC domain protein family, and its ability to disrupt microtubule-dependent vesicle transport relies on a catalytically active FIC domain (73, 81). The newly identified FIC domain protein Lpw_20091 shows little similarity to LegA8, suggesting that it might have a different host cell target and function.

TABLE 5. Putative novel Dot/Icm T4SS effectors and their homologues (>80% identity) in other *L. pneumophila* strains

| 130b effector | Homologue in strain: | | | | Domain/motif |
|---|---|---|---|---|---|
| | Philadelphia | Lens | Paris | Corby | |
| Lpw_00581 (LtpA) | | L*pl0057* | | L*pc0063*, L*pc3099* | Radial spoke head protein |
| Lpw_02251 | | | | | Calcineurin-like phosphoesterase domain |
| Lpw_02301 (LtpB) | | | | | Peptidase C58, Ankyrin repeat |
| Lpw_02381 (LtpC) | | | | | Ankyrin repeat, RasGEF domain |
| Lpw_03701 (LtpD) | | | | | |
| Lpw_04551 (LtpE) | | | | | |
| Lpw_16311 (LtpF) | | | | | Leucine-rich repeats |
| Lpw_20091 (LtpG) | | L*pl1931* | | L*pc1435* | FIC domain |
| Lpw_20341 (LtpH) | | | | | Ankyrin repeat |
| Lpw_21901 | | | | | Coiled coil |
| Lpw_25791 (LtpI) | | L*pl2297* | L*pp2417* | | |
| Lpw_25801 | | | L*pp2418* | L*pc2130* | Internal repeat |
| Lpw_25861[a] | | | | L*pc2106* | Host cell attachment protein |
| Lpw_26201 (LtpJ) | | L*pl2330* | | | Leucine-rich repeats |
| Lpw_28181 | | | | | Pentapeptide repeat |
| Lpw_28221 | | | | | NTPase (NACHT family) |

[a] Pseudogene in the genome assembly, but likely sequencing error.

The second new putative effector, Lpw_20341, is 130b strain specific and encoded next to the LegG1/PieG effector homologue Lpw_20351 (Lpg1975/6 in Philadelphia), which is found in all sequenced *L. pneumophila* strains (Fig. 5 and Table 4). Lpw_20341 has some similarity to the ankyrin repeat superfamily of proteins. Ankyrin repeats are predominantly eukaryotic protein-protein interaction motifs but are also found in 29 proteins of different *L. pneumophila* strains, 14 of which are proven Dot/Icm T4SS effectors (16, 33, 41, 73).

The genes encoding two additional novel ankyrin repeat proteins, Lpw_02301 and Lpw_02381, which have no homologues in the other sequenced *L. pneumophila* strains, are found in the ROD next to the *lvh2* locus in the 130b genome. Lpw_02301 has a predicted amino-terminal peptidase domain. Lpw_02381 is a large, 1,059-amino-acid protein that not only features 3 predicted ankyrin repeats in its N-teminal region, but also contains a predicted human Ras GTPase guanine nucleotide exchange factor (GEF) domain in its carboxyl terminus. The Ras subfamily of Ras GTPases is involved in the regulation of a wide variety of cellular processes, such as cell proliferation, adhesion, movement, division, secretion, and differentiation, which makes the members attractive targets for bacterial effector proteins (28, 64). The only other *L. pneumophila* protein for which a RasGEF domain has been identified is the effector protein LegG2, which is present in all the sequenced *L. pneumophila* strains (24).

The 130b genome encodes another ankyrin repeat protein, Lpw_22981, a homologue of which is found in Lens (Lpl2058). Genome comparison indicated that there is also a homologue of this gene in the Philadelphia-1 genome but that it has been fragmented by a deletion and multiple frameshift mutations. The predicted CDSs for the hypothetical proteins encoded by *lpg2128*, *lpg2129*, and *legA6* (*lpg2131*) together make up part of the full-length gene homologue. Fragments of an Lpw_22981 gene homologue are also present in Corby (*lpc1576*) and Paris (*lpp2068*). The fragmentation is a likely explanation as to why no translocation into host cells was observed for LegA6 (24).

Apart from the family of ankyrin repeat effectors, two other multitudinous effector families in *L. pneumophila* are formed by leucine-rich-repeat (LRR) and coiled-coil (CC) domain-containing effectors. LRRs are small sequence motifs that mediate protein-protein interactions of various proteins with diverse cellular functions. We identified two putative LRR effector proteins in 130b, Lpw_16311 and Lpw_26201. Lpw_16311 has no homologues in the other sequenced *L. pneumophila* strains, but the LRR domain has some similarity to other LRR-containing effectors, such as LegL7. Lpw_26201 is encoded next to the Dot/Icm effector Lem23 (Lpw_26191) in another previously identified hypervariable effector-rich genomic region located between *lpw_26101* and *lpw_26281* (*lpg2398* to *lpg2411*) (98). A homologue of Lpw_26201 is present in *L. pneumophila* Lens (Lpl2330), and it was shown that Lpl2330 is upregulated during the transmissive phase (52).

A new putative CC effector protein, Lpw_21901, which is found only in *L. pneumophila* 130b, is encoded in a ROD downstream of LGI2. CC domains induce protein-protein binding by homotypic interaction, and CC effector proteins, including LegC2, LegC7, and VipA, were shown to be involved in the modulation of vesicular trafficking (11, 87).

*L. pneumophila* 130b encodes several other putative strain-specific effector proteins scattered around the genome, as shown in Fig. 1. Lpw_03701 is encoded adjacent to a hypothetical protein, Lpw_03691, and together they represent an insertion in the 130b genome or a deletion in the other *L. pneumophila* genomes in an otherwise conserved region between the effectors LegG2 (Lpw_03651) and Ceg10 (Lpw_03741). The last 180 amino acids of this unique protein share 67% identity with the carboxyl terminus of an ankyrin repeat protein from *L. pneumophila* Paris (Lpp0356). However, the ankyrin repeats found in the amino terminus of Lpp0356 are not conserved in Lpw_03701. Instead, the amino terminus has similarity (32% identity) to the uncharacterized protein Lpw_27671, which has homologues in all five *L. pneumophila* strains. Another insertion in the 130b genome, Lpw_04551, is found in an otherwise conserved region between the proven effectors Lpg0365 (Lpw_04441) and SdhA (Lpw_04591). Lpw_04551 has some similarity to several *L. longbeachae* proteins, with the putative F-box protein Llb_3234 being the closest homologue (51% identity), and hypothetical proteins encoded in *L. pneumophila* Paris (Lpp1330; 50% identity) and the obligate intracellular amoebal pathogen *Legionella drancourtii* LLAP12 (46% identity).

Curiously, *L. pneumophila* 130b encodes a predicted nucleoside triphosphatase (NTPase) of the NACHT family (Lpw_28221) that has no homologues in other *L. pneumophila* strains. NACHT domains are found in animal, fungal, and bacterial proteins, and several NACHT domain-containing proteins, for example, Nod1/2 or the NAIPs, are involved in sensing bacterial pathogens and are important effector molecules of the innate immune response (8, 35). The NACHT domain protein NAIP5 is crucial for the activation of immune signaling and restriction of intracellular growth of *Legionella* in mouse macrophages (26, 95). It will be very interesting to analyze if *L. pneumophila* 130b might express a NACHT protein, which could enable the manipulation of innate immune recognition.

*L. pneumophila* 130b possesses several paralogues of genes encoding known effector proteins listed in Table 4, including a paralogue (*lpw_00591*) of *sdbB* (*lpw_27041*). The same set of two *sdbB* genes is also found in the same genomic locations in Lens (*lpl0058* and *lpl2402*) and Corby (*lpc0065* and *lpc1996*), but Philadelphia-1 and Paris each have only one of the two paralogous genes (*lpg2482* and *lpp2546*, respectively), as the *lpw_00591* homologue is absent from these strains.

Adjacent to the effector paralogue *sdbB2*, the 130b genome encodes another putative effector, Lpw_00581, which has some predicted structural similarity to a protozoan ciliary protein and is also found in *L. pneumophila* Lens and Corby.

**Ten novel 130b effectors are translocated into macrophages.** To investigate if some of the newly identified effector candidates were indeed substrates of the Dot/Icm T4SS, we used the fluorescence-based β-lactamase (TEM1) translocation assay to study effector translocation (17, 24). *L. pneumophila* 130b expressing N-terminal fusions of TEM1 to several putative effector proteins (see Figure S1 in the supplemental material) was used to infect Raw264.7 macrophages. Translocation of the fusion proteins into the host cells results in the cleavage of the TEM1 substrate CCF2, which is measured as a shift in its fluorescence emission wavelength. Our experiments demonstrated that TEM1 fusions of the effector protein candidates

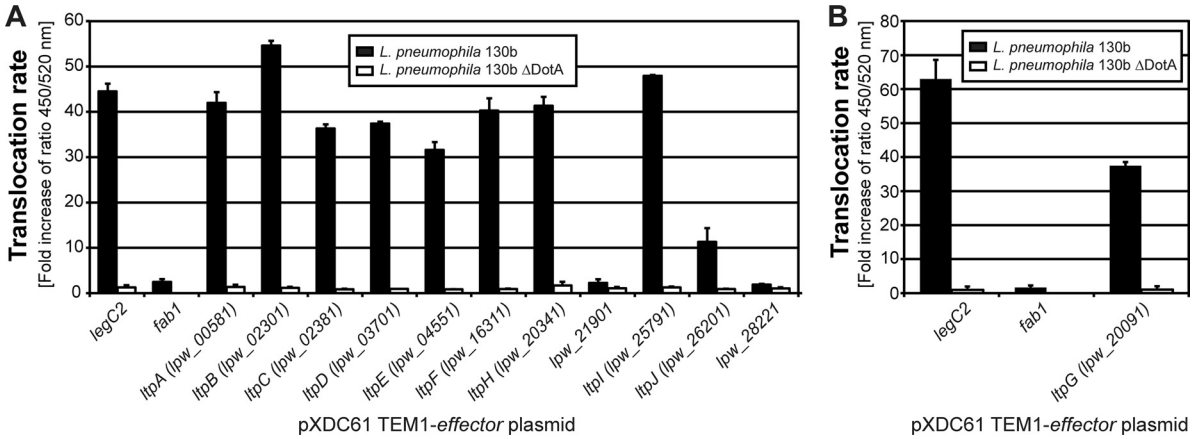


FIG. 6. *L. pneumophila* 130b translocates 10 novel Dot/Icm T4SS effector proteins into Raw264.7 macrophages. *L. pneumophila* 130b wild type (black bars) or the T4S-deficient *L. pneumophila* 130b ΔDotA mutant (white bars) harboring pXDC61 TEM1-effector gene constructs were grown for 21 h in the presence of 0.5 mM IPTG to induce expression of the TEM1-effector fusion proteins and used to infect macrophages (MOI, 40) (A) or grown for 16 h, and expression of TEM1-effector fusions was induced 5 h prior to and parallel to the infection (1 mM IPTG; MOI, 40) (B). The translocation of the TEM1 fusions of the proven Dot/Icm T4SS effector LegC2 (positive control), the new putative effectors, and the negative control Fab1 was measured using a Fluostar Optima plate reader (410-nm [10-nm band-pass] excitation and 450-nm and 520-nm emission filters). The translocation rate is expressed for each sample in relation to the emission ratio of uninfected cells. The fusion proteins of LegC2 and the *Legionella* translocated proteins LtpA to LtpJ, but not the negative control (Fab1) or the effector candidates Lpw_21901 and Lpw_28221, were translocated into host cells. The error bars represent standard deviations (SD). Similar results were obtained in three independent experiments.

Lpw_00581, Lpw_02301, Lpw_02381, Lpw_03701, Lpw_04551, Lpw_16311, Lpw_20341, Lpw_25791, and Lpw_26201 were translocated into Raw264.7 macrophages in a DotA-dependent manner, while TEM1-Lpw_21901 and TEM1-Lpw_28221 were not translocated into the host cells (Fig. 6A). Interestingly, the induction of expression of TEM1-Lpw_20091 in *Legionella* overnight cultures led to a strong growth defect (data not shown). To circumvent this problem, we grew the bacterial cultures without induction for 16 h and then induced expression of TEM1-Lpw_20091 and the positive control TEM1-LegC2 for 5 h and during the first hour of infection (Fig. 6B) or only parallel to the infection (data not shown). Under these conditions, we observed translocation of TEM1-Lpw_20091 by wild-type bacteria to levels comparable to those of TEM1-effector fusion proteins.

Altogether, we identified 10 novel T4SS effector proteins, which we called *Legionella* translocated proteins LtpA to LtpJ (Table 5), that are not present in *L. pneumophila* Philadelphia-1. Among them, the FIC domain effector LtpG (Lpw_20091) showed a particular phenotype, as induction of protein expression in liquid cultures resulted in a strong growth defect. No bacterial growth defects upon heterologous expression of the FIC domain effector AnkX were reported. It is still not known if the growth defect depends on the enzymatic activity of the FIC domain or how expression and translocation of endogenous LtpG are regulated during infection.

**Prevalence of the novel effector proteins in a collection of *Legionella* isolates.** The newly identified T4SS effector proteins were selected because they were absent from *L. pneumophila* Philadelphia. To determine the prevalence of these effectors outside the sequenced prototypic strains, we performed a PCR screen of a large set of environmental and clinical *Legionella* isolates (summarized in Fig. 7; individual results for each strain are shown in Table S2 in the supplemental material). The results for representative PCR-negative strains were confirmed by Southern blotting (data not shown). The set included 87 environmental and clinical *L. pneumophila* isolates from different serogroups from Malaysia, Australia, and the United Kingdom. Based on this screen, the novel effectors could be divided into two groups (Fig. 7). Group 1 comprised rare effector genes, namely, *ltpB*, *ltpC*, *ltpE*, *ltpF*, and *ltpJ*, which were detected in less than 15% of the analyzed strains. In contrast, members of the second group, *ltpA*, *ltpD*, *ltpG*, *ltpH*, and *ltpI*, showed higher prevalences of 34% to 62% in the PCR screen. The prevalences of three *L. pneumophila* Lens homologues of the newly identified effectors *ltpA*, *ltpG*, and *ltpJ* had already been investigated in a seminal multigenome DNA array screen including more than 200 *L. pneumophila* and non-

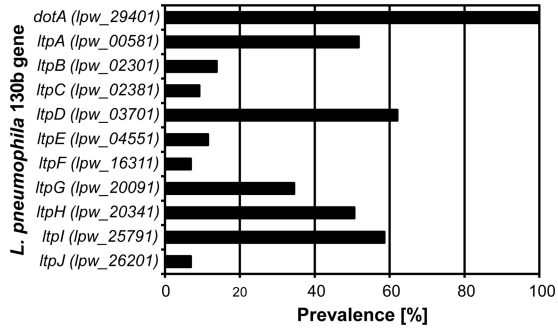


FIG. 7. Distribution of the 10 novel Dot/Icm T4SS effectors in 87 *L. pneumophila* isolates. Genomic DNA from 87 clinical and environmental *L. pneumophila* isolates was analyzed for the presence of homologues of the novel *L. pneumophila* 130b effectors *ltpA* to *ltpJ* by PCR. The *dotA* gene was included in the screen as a positive control for the presence of the Dot/Icm T4SSB. The new effectors fall in two classes. Class 1 is formed by the rare effectors *ltpB*, *ltpC*, *ltpE*, *ltpF*, and *ltpJ*, which were detected in less than 15% of the analyzed strains. Class 2 contains the more prevalent effectors *ltpA*, *ltpD*, *ltpG*, *ltpH*, and *ltpI*, which were found in 34% to 62% of the isolates.

*pneumophila* strains (15). In this screen, the gene *lpl0057* (an *ltpA* homologue) was found in 36%, *lpl1931* (an *ltpG* homologue) in 40%, and *lpl2330* (an *ltpJ* homologue) in 8% of all tested strains. These numbers show good correlation with our data (*ltpA*, 52%; *ltpG*, 34%; and *ltpJ*, 7%) and support the reliability of the PCR screen. Our screen revealed no correlation between the presence of the novel effectors and the clinical or environmental origin of the strains. However, there is no scientific evidence that these environmental *L. pneumophila* strains have less potential to cause Legionnaires' disease in humans than clinical strains. The results of our screen suggest that there is a prominent group of effectors present in 30 to 60% of the population, which has to be considered when Dot/Icm T4SS-dependent phenotypes are compared between different strains or generalized for the species *L. pneumophila*.

**Conclusions.** The genome of *L. pneumophila* strain Alcoy (accession number CP001828) became available just prior to the submission of the manuscript, and comparative genome analysis confirmed that the Dot/Icm T4SSB is conserved among all the *L. pneumophila* strains, although this is not the case for the T4SSA systems. Strains Paris, Lens, and Philadelphia-1 do not have the Tra/Trb system, and Corby does not have an Lvh region; however, 130b has both T4SSA systems. Both regions were also recently reported to be present in the genome of Alcoy (23). However, we report here the first identification of two distinct Lvh regions within one strain, and it will be interesting to see if both regions are functional and contribute to virulence. Analysis of the *L. pneumophila* Alcoy genome revealed that all 107 of the putative core set of *L. pneumophila* Dot/Icm T4SS effectors identified in this study are present. In addition, the redefined hypervariable region we describe here correlates with a synonymous region in Alcoy (*lpa_02788* to *lpa_02905*). No new Dot/Icm T4SS effectors have been described for strain Alcoy. Of the 10 newly identified and validated T4SS effectors, only LtpA and LtpE possess homologues with moderate similarity in the strain, underlining the high plasticity of the effector pool between *L. pneumophila* strains. The characterization of the mode of action of the novel effector proteins will reveal if and how these effectors contribute to strain-to-strain variation in virulence.

## ACKNOWLEDGMENTS

We are indebted to Howard Shuman and Xavier Charpentier for the generous gift of pXDC61 and the control plasmids for the TEM1 translocation assay. We thank Anthony Davidson for his assistance with the PCR screen. We are indebted to Stacey Yong for providing Malaysian *L. pneumophila* isolates for screening.

E.H. is an Australian Research Council (ARC) Future Fellow, and S.A.B. is an ARC Research Fellow. The Pathogen genomics group is funded by The Wellcome Trust. This work was supported by the National Health and Medical Research Council (NHMRC) of Australia, the Wellcome Trust, and the Medical Research Council (MRC).

## REFERENCES

1. **Alli, O. A., S. Zink, N. K. von Lackum, and Y. Abu-Kwaik.** 2003. Comparative assessment of virulence traits in *Legionella* spp. Microbiology **149:**631–641.
2. **Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman.** 1990. Basic local alignment search tool. J. Mol. Biol. **215:**403–410.
3. **Alvarez-Martinez, C. E., and P. J. Christie.** 2009. Biological diversity of prokaryotic type IV secretion systems. Microbiol. Mol. Biol. Rev. **73:**775–808.
4. **Asare, R., M. Santic, I. Gobin, M. Doric, J. Suttles, J. E. Graham, C. D. Price, and Y. Abu Kwaik.** 2007. Genetic susceptibility and caspase activation in mouse and human macrophages are distinct for *Legionella longbeachae* and *L. pneumophila*. Infect. Immun. **75:**1933–1945.
5. **Backert, S., and T. F. Meyer.** 2006. Type IV secretion systems and their effectors in bacterial pathogenesis. Curr. Opin. Microbiol. **9:**207–217.
6. **Bandyopadhyay, P., S. Liu, C. B. Gabbai, Z. Venitelli, and H. M. Steinman.** 2007. Environmental mimics and the Lvh type IVA secretion system contribute to virulence-related phenotypes of *Legionella pneumophila*. Infect. Immun. **75:**723–735.
7. **Brüggemann, H., C. Cazalet, and C. Buchrieser.** 2006. Adaptation of *Legionella pneumophila* to the host environment: role of protein secretion, effectors and eukaryotic-like proteins. Curr. Opin. Microbiol. **9:**86–94.
8. **Bryant, C., and K. A. Fitzgerald.** 2009. Molecular mechanisms involved in inflammasome activation. Trends Cell Biol. **19:**455–464.
9. **Burstein, D., T. Zusman, E. Degtyar, R. Viner, G. Segal, and T. Pupko.** 2009. Genome-scale identification of *Legionella pneumophila* effectors using a machine learning approach. PLoS Pathog. **5:**e1000508.
10. **Cambronne, E. D., and C. R. Roy.** 2006. Recognition and delivery of effector proteins into eukaryotic cells by bacterial secretion systems. Traffic **7:**929–939.
11. **Campodonico, E. M., L. Chesnel, and C. R. Roy.** 2005. A yeast genetic system for the identification and characterization of substrate proteins transferred into host cells by the *Legionella pneumophila* Dot/Icm system. Mol. Microbiol. **56:**918–933.
12. **Carver, T., N. Thomson, A. Bleasby, M. Berriman, and J. Parkhill.** 2009. DNAPlotter: circular and linear interactive genome visualization. Bioinformatics **25:**119–120.
13. **Carver, T. J., K. M. Rutherford, M. Berriman, M. A. Rajandream, B. G. Barrell, and J. Parkhill.** 2005. ACT: the Artemis Comparison Tool. Bioinformatics **21:**3422–3423.
14. **Cazalet, C., L. Gomez-Valero, C. Rusniok, M. Lomma, D. Dervins-Ravault, H. J. Newton, F. M. Sansom, S. Jarraud, N. Zidane, L. Ma, C. Bouchier, J. Etienne, E. L. Hartland, and C. Buchrieser.** 2010. Analysis of the *Legionella longbeachae* genome and transcriptome uncovers unique strategies to cause Legionnaires' disease. PLoS Genet. **6:**e1000851.
15. **Cazalet, C., S. Jarraud, Y. Ghavi-Helm, F. Kunst, P. Glaser, J. Etienne, and C. Buchrieser.** 2008. Multigenome analysis identifies a worldwide distributed epidemic *Legionella pneumophila* clone that emerged within a highly diverse species. Genome Res. **18:**431–441.
16. **Cazalet, C., C. Rusniok, H. Bruggemann, N. Zidane, A. Magnier, L. Ma, M. Tichit, S. Jarraud, C. Bouchier, F. Vandenesch, F. Kunst, J. Etienne, P. Glaser, and C. Buchrieser.** 2004. Evidence in the *Legionella pneumophila* genome for exploitation of host cell functions and high genome plasticity. Nat. Genet. **36:**1165–1173.
17. **Charpentier, X., J. E. Gabay, M. Reyes, J. W. Zhu, A. Weiss, and H. A. Shuman.** 2009. Chemical genetics reveals bacterial and host cell functions critical for type IV effector translocation by *Legionella pneumophila*. PLoS Pathog. **5:**e1000501.
18. **Chien, M., I. Morozova, S. Shi, H. Sheng, J. Chen, S. M. Gomez, G. Asamani, K. Hill, J. Nuara, M. Feder, J. Rineer, J. J. Greenberg, V. Steshenko, S. H. Park, B. Zhao, E. Teplitskaya, J. R. Edwards, S. Pampou, A. Georghiou, I. C. Chou, W. Iannuccilli, M. E. Ulz, D. H. Kim, A. Geringer-Sameth, C. Goldsberry, P. Morozov, S. G. Fischer, G. Segal, X. Qu, A. Rzhetsky, P. Zhang, E. Cayanis, P. J. De Jong, J. Ju, S. Kalachikov, H. A. Shuman, and J. J. Russo.** 2004. The genomic sequence of the accidental pathogen *Legionella pneumophila*. Science **305:**1966–1968.
19. **Christie, P. J., K. Atmakuri, V. Krishnamoorthy, S. Jakubowski, and E. Cascales.** 2005. Biogenesis, architecture, and function of bacterial type IV secretion systems. Annu. Rev. Microbiol. **59:**451–485.
20. **Cianciotto, N. P.** 2009. Many substrates and functions of type II secretion: lessons learned from *Legionella pneumophila*. Future Microbiol. **4:**797–805.
21. **Cirillo, S. L., J. Lum, and J. D. Cirillo.** 2000. Identification of novel loci involved in entry by *Legionella pneumophila*. Microbiology **146:**1345–1359.
22. **Coburn, B., I. Sekirov, and B. B. Finlay.** 2007. Type III secretion systems and disease. Clin. Microbiol. Rev. **20:**535–549.
23. **D'Auria, G., N. Jimenez-Hernandez, F. Peris-Bondia, A. Moya, and A. Latorre.** 2010. *Legionella pneumophila* pangenome reveals strain-specific virulence factors. BMC Genomics **11:**181.
24. **de Felipe, K. S., R. T. Glover, X. Charpentier, O. R. Anderson, M. Reyes, C. D. Pericone, and H. A. Shuman.** 2008. *Legionella* eukaryotic-like type IV substrates interfere with organelle trafficking. PLoS Pathog. **4:**e1000117.
25. **de Felipe, K. S., S. Pampou, O. S. Jovanovic, C. D. Pericone, S. F. Ye, S. Kalachikov, and H. A. Shuman.** 2005. Evidence for acquisition of *Legionella* type IV secretion substrates via interdomain horizontal gene transfer. J. Bacteriol. **187:**7716–7726.
26. **Diez, E., S. H. Lee, S. Gauthier, Z. Yaraghi, M. Tremblay, S. Vidal, and P. Gros.** 2003. Birc1e is the gene within the Lgn1 locus associated with resistance to *Legionella pneumophila*. Nat. Genet. **33:**55–60.
27. **Doléans-Jordheim, A., M. Akermi, C. Ginevra, C. Cazalet, E. Kay, D. Schneider, C. Buchrieser, D. Atlan, F. Vandenesch, J. Etienne, and S. Jarraud.** 2006. Growth-phase-dependent mobility of the lvh-encoding region in *Legionella pneumophila* strain Paris. Microbiology **152:**3561–3568.
28. **Drosten, M., A. Dhawahir, E. Y. Sum, J. Urosevic, C. G. Lechuga, L. M.**

**Esteban, E. Castellano, C. Guerra, E. Santos, and M. Barbacid.** 2010. Genetic analysis of Ras signalling pathways in cell proliferation, migration and survival. EMBO J. **29:**1091–1104.
29. **Edelstein, P. H.** 1986. Control of *Legionella* in hospitals. J. Hosp. Infect. **8:**109–115.
30. **Engleberg, N. C., D. J. Drutz, and B. I. Eisenstein.** 1984. Cloning and expression of *Legionella pneumophila* antigens in *Escherichia coli*. Infect. Immun. **44:**222–227.
31. **Ensminger, A. W., and R. R. Isberg.** 2009. *Legionella pneumophila* Dot/Icm translocated substrates: a sum of parts. Curr. Opin. Microbiol. **12:**67–73.
32. **Finn, R. D., J. Tate, J. Mistry, P. C. Coggill, S. J. Sammut, H. R. Hotz, G. Ceric, K. Forslund, S. R. Eddy, E. L. Sonnhammer, and A. Bateman.** 2008. The Pfam protein families database. Nucleic Acids Res. **36:**D281–D288.
33. **Franco, I. S., H. A. Shuman, and X. Charpentier.** 2009. The perplexing functions and surprising origins of *Legionella pneumophila* type IV secretion effectors. Cell. Microbiol. **11:**1435–1443.
34. **Fraser, D. W., T. R. Tsai, W. Orenstein, W. E. Parkin, H. J. Beecham, R. G. Sharrar, J. Harris, G. F. Mallison, S. M. Martin, J. E. McDade, C. C. Shepard, and P. S. Brachman.** 1977. Legionnaires' disease: description of an epidemic of pneumonia. N. Engl. J. Med. **297:**1189–1197.
35. **Fritz, J. H., R. L. Ferrero, D. J. Philpott, and S. E. Girardin.** 2006. Nod-like proteins in immunity, inflammation and disease. Nat. Immunol. **7:**1250–1257.
36. **Fronzes, R., P. J. Christie, and G. Waksman.** 2009. The structural biology of type IV secretion systems. Nat. Rev. Microbiol. **7:**703–714.
37. **Glöckner, G., C. Albert-Weissenberger, E. Weinmann, S. Jacobi, E. Schunder, M. Steinert, J. Hacker, and K. Heuner.** 2008. Identification and characterization of a new conjugation/type IVA secretion system (trb/tra) of *Legionella pneumophila* Corby localized on two mobile genomic islands. Int. J. Med. Microbiol. **298:**411–428.
38. **Gomez-Valero, L., C. Rusniok, and C. Buchrieser.** 2009. *Legionella pneumophila*: population genetics, phylogeny and genomics. Infect. Genet. Evol. **9:**727–739.
39. **Gordon, D., C. Abajian, and P. Green.** 1998. Consed: a graphical tool for sequence finishing. Genome Res. **8:**195–202.
40. **Guindon, S., and O. Gascuel.** 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst. Biol. **52:**696–704.
41. **Habyarimana, F., S. Al-Khodor, A. Kalia, J. E. Graham, C. T. Price, M. T. Garcia, and Y. A. Kwaik.** 2008. Role for the Ankyrin eukaryotic-like genes of *Legionella pneumophila* in parasitism of protozoan hosts and human macrophages. Environ. Microbiol. **10:**1460–1474.
42. **Harrison, T. G., B. Afshar, N. Doshi, N. K. Fry, and J. V. Lee.** 2009. Distribution of *Legionella pneumophila* serogroups, monoclonal antibody subgroups and DNA sequence types in recent clinical and environmental isolates from England and Wales (2000–2008). Eur. J. Clin. Microbiol. Infect. Dis. **28:**781–791.
43. **Helbig, J. H., S. Bernander, M. Castellani Pastoris, J. Etienne, V. Gaia, S. Lauwers, D. Lindsay, P. C. Luck, T. Marques, S. Mentula, M. F. Peeters, C. Pelaz, M. Struelens, S. A. Uldum, G. Wewalka, and T. G. Harrison.** 2002. Pan-European study on culture-proven Legionnaires' disease: distribution of *Legionella pneumophila* serogroups and monoclonal subgroups. Eur. J. Clin. Microbiol. Infect. Dis. **21:**710–716.
44. **Hilbi, H., G. Segal, and H. A. Shuman.** 2001. Icm/dot-dependent upregulation of phagocytosis by *Legionella pneumophila*. Mol. Microbiol. **42:**603–617.
45. **Hilbi, H., S. S. Weber, C. Ragaz, Y. Nyfeler, and S. Urwyler.** 2007. Environmental predators as models for bacterial pathogenesis. Environ. Microbiol. **9:**563–575.
46. **Horwitz, M. A.** 1983. Formation of a novel phagosome by the Legionnaires' disease bacterium (*Legionella pneumophila*) in human monocytes. J. Exp. Med. **158:**1319–1331.
47. **Horwitz, M. A., and F. R. Maxfield.** 1984. *Legionella pneumophila* inhibits acidification of its phagosome in human monocytes. J. Cell Biol. **99:**1936–1943.
48. **Isberg, R. R., T. J. O'Connor, and M. Heidtman.** 2009. The *Legionella pneumophila* replication vacuole: making a cozy niche inside host cells. Nat. Rev. Microbiol. **7:**13–24.
49. **Jacobi, S., and K. Heuner.** 2003. Description of a putative type I secretion system in *Legionella pneumophila*. Int. J. Med. Microbiol. **293:**349–358.
50. **Juhas, M., D. W. Crook, I. D. Dimopoulou, G. Lunter, R. M. Harding, D. J. Ferguson, and D. W. Hood.** 2007. Novel type IV secretion system involved in propagation of genomic islands. J. Bacteriol. **189:**761–771.
51. **Juhas, M., D. W. Crook, and D. W. Hood.** 2008. Type IV secretion systems: tools of bacterial horizontal gene transfer and virulence. Cell. Microbiol. **10:**2377–2386.
52. **Jules, M., and C. Buchrieser.** 2007. *Legionella pneumophila* adaptation to intracellular life and the host response: clues from genomics and transcriptomics. FEBS Lett. **581:**2829–2838.
53. **Kagan, J. C., and C. R. Roy.** 2002. *Legionella* phagosomes intercept vesicular traffic from endoplasmic reticulum exit sites. Nat. Cell Biol. **4:**945–954.
54. **Kinch, L. N., M. L. Yarbrough, K. Orth, and N. V. Grishin.** 2009. Fido, a novel AMPylation domain common to fic, doc, and AvrB. PLoS One **4:**e5818.
55. **Komano, T., R. Utsumi, and M. Kawamukai.** 1991. Functional analysis of the fic gene involved in regulation of cell division. Res. Microbiol. **142:**269–277.
56. **Kozak, N. A., M. Buss, C. E. Lucas, M. Frace, D. Govil, T. Travis, M. Olsen-Rasmussen, R. F. Benson, and B. S. Fields.** 2010. Virulence factors encoded by *Legionella longbeachae* identified on the basis of the genome sequence analysis of clinical isolate D-4968. J. Bacteriol. **192:**1030–1044.
57. **Kubori, T., A. Hyakutake, and H. Nagai.** 2008. *Legionella* translocates an E3 ubiquitin ligase that has multiple U-boxes with distinct functions. Mol. Microbiol. **67:**1307–1319.
58. **Lamoth, F., and G. Greub.** 2009. Amoebal pathogens as emerging causal agents of pneumonia. FEMS Microbiol. Rev. 10.1111/j.1574-6976.2009.00207.x.
59. **Larkin, M. A., G. Blackshields, N. P. Brown, R. Chenna, P. A. McGettigan, H. McWilliam, F. Valentin, I. M. Wallace, A. Wilm, R. Lopez, J. D. Thompson, T. J. Gibson, and D. G. Higgins.** 2007. Clustal W and Clustal X version 2.0. Bioinformatics **23:**2947–2948.
60. **Lehnherr, H., E. Maguin, S. Jafri, and M. B. Yarmolinsky.** 1993. Plasmid addiction genes of bacteriophage P1: doc, which causes cell death on curing of prophage, and phd, which prevents host death when prophage is retained. J. Mol. Biol. **233:**414–428.
61. **Letunic, I., T. Doerks, and P. Bork.** 2009. SMART 6: recent updates and new developments. Nucleic Acids Res. **37:**D229–D232.
62. **Lomma, M., D. Dervins-Ravault, M. Rolando, T. Nora, H. J. Newton, F. M. Samson, T. Sahr, L. Gomez-Valero, M. Jules, E. L. Hartland, and C. Buchrieser.** 2010. The *Legionella pneumophila* F-box protein Lpp2082 (AnkB) modulates ubiquitination of the host protein parvin B and promotes intracellular replication. Cell. Microbiol. 10.1111/j.1462-5822.2010.01467.x.
63. **Luo, Z. Q., and R. R. Isberg.** 2004. Multiple substrates of the *Legionella pneumophila* Dot/Icm system identified by interbacterial protein transfer. Proc. Natl. Acad. Sci. U. S. A. **101:**841–846.
64. **Malumbres, M., and M. Barbacid.** 2003. RAS oncogenes: the first 30 years. Nat. Rev. Cancer **3:**459–465.
65. **Marchler-Bauer, A., J. B. Anderson, F. Chitsaz, M. K. Derbyshire, C. DeWeese-Scott, J. H. Fong, L. Y. Geer, R. C. Geer, N. R. Gonzales, M. Gwadz, S. He, D. I. Hurwitz, J. D. Jackson, Z. Ke, C. J. Lanczycki, C. A. Liebert, C. Liu, F. Lu, S. Lu, G. H. Marchler, M. Mullokandov, J. S. Song, A. Tasneem, N. Thanki, R. A. Yamashita, D. Zhang, N. Zhang, and S. H. Bryant.** 2009. CDD: specific functional annotation with the Conserved Domain Database. Nucleic Acids Res. **37:**D205–D210.
66. **McDade, J. E., C. C. Shepard, D. W. Fraser, T. R. Tsai, M. A. Redus, and W. R. Dowdle.** 1977. Legionnaires' disease: isolation of a bacterium and demonstration of its role in other respiratory disease. N. Engl. J. Med. **297:**1197–1203.
67. **Mohd-Zain, Z., S. L. Turner, A. M. Cerdeno-Tarraga, A. K. Lilley, T. J. Inzana, A. J. Duncan, R. M. Harding, D. W. Hood, T. E. Peto, and D. W. Crook.** 2004. Transferable antibiotic resistance elements in *Haemophilus influenzae* share a common evolutionary origin with a diverse family of syntenic genomic islands. J. Bacteriol. **186:**8114–8122.
68. **Molmeret, M., M. Horn, M. Wagner, M. Santic, and Y. Abu Kwaik.** 2005. Amoebae as training grounds for intracellular bacterial pathogens. Appl. Environ. Microbiol. **71:**20–28.
69. **Morozova, I., X. Qu, S. Shi, G. Asamani, J. E. Greenberg, H. A. Shuman, and J. J. Russo.** 2004. Comparative sequence analysis of the icm/dot genes in *Legionella*. Plasmid **51:**127–147.
70. **Newton, H. J., D. K. Ang, I. R. van Driel, and E. L. Hartland.** 2010. Molecular pathogenesis of infections caused by *Legionella pneumophila*. Clin. Microbiol. Rev. **23:**274–298.
71. **Newton, H. J., F. M. Sansom, V. Bennett-Wood, and E. L. Hartland.** 2006. Identification of *Legionella pneumophila*-specific genes by genomic subtractive hybridization with *Legionella micdadei* and identification of *lpnE*, a gene required for efficient host cell entry. Infect. Immun. **74:**1683–1691.
72. **Ninio, S., J. Celli, and C. R. Roy.** 2009. A Legionella pneumophila effector protein encoded in a region of genomic plasticity binds to Dot/Icm-modified vacuoles. PLoS Pathog. **5:**e1000278.
73. **Pan, X., A. Luhrmann, A. Satoh, M. A. Laskowski-Arce, and C. R. Roy.** 2008. Ankyrin repeat proteins comprise a diverse family of bacterial type IV effectors. Science **320:**1651–1654.
74. **Posada, D., and K. A. Crandall.** 1998. MODELTEST: testing the model of DNA substitution. Bioinformatics **14:**817–818.
75. **Price, C. T., S. Al-Khodor, T. Al-Quadan, M. Santic, F. Habyarimana, A. Kalia, and Y. A. Kwaik.** 2009. Molecular mimicry by an F-box effector of *Legionella pneumophila* hijacks a conserved polyubiquitination machinery within macrophages and protozoa. PLoS Pathog. **5:**e1000704.
76. **Pukatzki, S., S. B. McAuley, and S. T. Miyata.** 2009. The type VI secretion system: translocation of effectors and effector-domains. Curr. Opin. Microbiol. **12:**11–17.
77. **Ratzow, S., V. Gaia, J. H. Helbig, N. K. Fry, and P. C. Luck.** 2007. Addition of neuA, the gene encoding N-acylneuraminate cytidylyl transferase, increases the discriminatory ability of the consensus sequence-based scheme for typing *Legionella pneumophila* serogroup 1 strains. J. Clin. Microbiol. **45:**1965–1968.
78. **Ridenour, D. A., S. L. Cirillo, S. Feng, M. M. Samrakandi, and J. D. Cirillo.** 2003. Identification of a gene that affects the efficiency of host cell infection

by *Legionella pneumophila* in a temperature-dependent fashion. Infect. Immun. **71:**6256–6263.
79. **Rossier, O., and N. P. Cianciotto.** 2001. Type II protein secretion is a subset of the PilD-dependent processes that facilitate intracellular infection by *Legionella pneumophila*. Infect. Immun. **69:**2092–2098.
80. **Rowbotham, T. J.** 1980. Preliminary report on the pathogenicity of *Legionella pneumophila* for freshwater and soil amoebae. J. Clin. Pathol. **33:**1179–1183.
81. **Roy, C. R., and S. Mukherjee.** 2009. Bacterial FIC proteins AMP up infection. Sci. Signal. **2:**pe14.
82. **Rutherford, K., J. Parkhill, J. Crook, T. Horsnell, P. Rice, M. A. Rajandream, and B. Barrell.** 2000. Artemis: sequence visualization and annotation. Bioinformatics **16:**944–945.
83. **Samrakandi, M. M., S. L. Cirillo, D. A. Ridenour, L. E. Bermudez, and J. D. Cirillo.** 2002. Genetic and phenotypic differences between *Legionella pneumophila* strains. J. Clin. Microbiol. **40:**1352–1362.
84. **Sansom, F. M., H. J. Newton, S. Crikis, N. P. Cianciotto, P. J. Cowan, A. J. d'Apice, and E. L. Hartland.** 2007. A bacterial ecto-triphosphate diphosphohydrolase similar to human CD39 is essential for intracellular multiplication of *Legionella pneumophila*. Cell. Microbiol. **9:**1922–1935.
85. **Segal, G., M. Purcell, and H. A. Shuman.** 1998. Host cell killing and bacterial conjugation require overlapping sets of genes within a 22-kb region of the *Legionella pneumophila* genome. Proc. Natl. Acad. Sci. U. S. A. **95:**1669–1674.
86. **Segal, G., J. J. Russo, and H. A. Shuman.** 1999. Relationships between a new type IV secretion system and the icm/dot virulence system of *Legionella pneumophila*. Mol. Microbiol. **34:**799–809.
87. **Shohdy, N., J. A. Efe, S. D. Emr, and H. A. Shuman.** 2005. Pathogen effector protein screening in yeast identifies *Legionella* factors that interfere with membrane trafficking. Proc. Natl. Acad. Sci. U. S. A. **102:**4866–4871.
88. **Steinert, M., K. Heuner, C. Buchrieser, C. Albert-Weissenberger, and G. Glockner.** 2007. *Legionella* pathogenicity: genome structure, regulatory networks and the host cell response. Int. J. Med. Microbiol. **297:**577–587.
89. **Swanson, M. S., and R. R. Isberg.** 1995. Association of *Legionella pneumophila* with the macrophage endoplasmic reticulum. Infect. Immun. **63:**3609–3620.
90. **Thomson, N. R., D. J. Clayton, D. Windhorst, G. Vernikos, S. Davidson, C. Churcher, M. A. Quail, M. Stevens, M. A. Jones, M. Watson, A. Barron, A. Layton, D. Pickard, R. A. Kingsley, A. Bignell, L. Clark, B. Harris, D. Ormond, Z. Abdellah, K. Brooks, I. Cherevach, T. Chillingworth, J. Woodward, H. Norberczak, A. Lord, C. Arrowsmith, K. Jagels, S. Moule, K. Mungall, M. Sanders, S. Whitehead, J. A. Chabalgoity, D. Maskell, T. Humphrey, M. Roberts, P. A. Barrow, G. Dougan, and J. Parkhill.** 2008. Comparative genome analysis of *Salmonella* Enteritidis PT4 and *Salmonella* Gallinarum 287/91 provides insights into evolutionary and host adaptation pathways. Genome Res. **18:**1624–1637.
91. **Vogel, J. P., H. L. Andrews, S. K. Wong, and R. R. Isberg.** 1998. Conjugative transfer by the virulence system of *Legionella pneumophila*. Science **279:**873–876.
92. **Weber, S. S., C. Ragaz, K. Reus, Y. Nyfeler, and H. Hilbi.** 2006. *Legionella pneumophila* exploits PI(4)P to anchor secreted effector proteins to the replicative vacuole. PLoS Pathog. **2:**e46.
93. **Wilkinson, P., K. Paszkiewicz, A. Moorhouse, J. M. Szubert, S. Beatson, J. Gerrard, N. R. Waterfield, and R. H. Ffrench-Constant.** 2010. New plasmids and putative virulence factors from the draft genome of an Australian clinical isolate of *Photorhabdus asymbiotica*. FEMS Microbiol. Lett. **309:**136–143.
94. **Worby, C. A., S. Mattoo, R. P. Kruger, L. B. Corbeil, A. Koller, J. C. Mendez, B. Zekarias, C. Lazar, and J. E. Dixon.** 2009. The FIC domain: regulation of cell signaling by adenylylation. Mol. Cell **34:**93–103.
95. **Wright, E. K., S. A. Goodart, J. D. Growney, V. Hadinoto, M. G. Endrizzi, E. M. Long, K. Sadigh, A. L. Abney, I. Bernstein-Hanley, and W. F. Dietrich.** 2003. Naip5 affects host susceptibility to the intracellular pathogen *Legionella pneumophila*. Curr. Biol. **13:**27–36.
96. **Wu, M., and J. A. Eisen.** 2008. A simple, fast, and accurate method of phylogenomic inference. Genome Biol. **9:**R151.
97. **Yarbrough, M. L., Y. Li, L. N. Kinch, N. V. Grishin, H. L. Ball, and K. Orth.** 2009. AMPylation of Rho GTPases by *Vibrio* VopS disrupts effector binding and downstream signaling. Science **323:**269–272.
98. **Zusman, T., E. Degtyar, and G. Segal.** 2008. Identification of a hypervariable region containing new *Legionella pneumophila* Icm/Dot translocated substrates by using the conserved icmQ regulatory signature. Infect. Immun. **76:**4581–4591.